DOCUMENT RESUME

ED 185 193 UD 020 455

AUTHOR Loewen, James W.; And Others
TITLE Social Science for a Change.
INSTITUTION Catholic Univ. of America, Washington, D.C. Center for National Policy Review.
PUB DATE 79
NOTE 41p.; Not available in paper copy due to print color of original document
JOURNAL CIT Clearinghouse for Civil Rights Research; v7 n3-4 Fall-Win 1979

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Civil Rights; *Court Litigation; Elections; Employment Practices; Housing; Minority Groups; Promotion (Occupational); Racial Segregation; Schools; *Social Discrimination; *Social Science Research; *Statistics; Voting

ABSTRACT

This publication is a collection of social science research and statistical techniques that can help win civil rights court cases. Articles included are: (1) "A Lawyer Looks at Social Science in the Courts"; (2) "How Social Scientists and Lawyers Can Work Together"; (3) "The Nature of Statistics and Research as Used in Civil Rights Litigation"; (4) "The Sign Test for Comparing Hiring/Promotion Records for Possible Discrimination"; (5) "The T Test for the Difference of Two Proportions: Analyzing Juries and Venires for Discrimination"; (6) "Selecting a Defensible Sample and Using the T Test for the Difference of Two Means in Municiple Services Discrimination Cases"; (7) "Using the Index of Dissimilarity to Determine the Extent of School or Residential Segregation"; (8) "The Runs Test for Determining if Minorities are Being Clustered within a Dormitory or Housing Project"; and (9) "Using Correlation, Overlapping Percentages, Regression, and Census Data to Analyze Election Returns for Racial Bloc Voting." References are appended. (MK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

MAR 1 2 1980

Clearinghouse for Civil Rights Research

VOLUME VII # 3-4
AUTUMN-WINTER, 1979

A publication of the Center for National Policy Review

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. Loewen
Center for National Policy Review
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SOCIAL SCIENCE FOR A CHANGE

A collection of social science research and statistical techniques that can help win civil rights cases.

2 INTRODUCTION: A LAWYER LOOKS AT SOCIAL SCIENCE IN THE COURTS.

5 HOW SOCIAL SCIENTISTS AND LAWYERS CAN WORK TOGETHER.

10 THE NATURE OF STATISTICS AND RESEARCH AS USED IN CIVIL RIGHTS LITIGATION.

16 THE SIGN TEST FOR COMPARING HIRING/PROMOTION RECORDS FOR POSSIBLE DISCRIMINATION.

18 THE T TEST FOR THE DIFFERENCE OF TWO PROPORTIONS: ANALYZING JURIES AND VENIRES FOR DISCRIMINATION.

22 SELECTING A DEFENSIBLE SAMPLE AND USING THE T TEST FOR THE DIFFERENCE OF TWO MEANS IN MUNICIPAL SERVICES DISCRIMINATION CASES.

27 USING THE INDEX OF DISSIMILARITY TO DETERMINE THE EXTENT OF SCHOOL OR RESIDENTIAL SEGREGATION.

29 THE RUNS TEST FOR DETERMINING IF MINORITIES ARE BEING CLUSTERED WITHIN A DORMITORY OR HOUSING PROJECT.

30 USING CORRELATION, OVERLAPPING PERCENTAGES, REGRESSION, AND CENSUS DATA TO ANALYZE ELECTION RETURNS FOR RACIAL BLOC VOTING.

35 REFERENCES.

38 DEPARTMENTS.
Contributors
Communications to the CLEARINGHOUSE
Seeking Social Science

The introduction is by William L. Taylor; the articles are by James W. Loewen; Kathleen H. Courtney is co-author of the last article and the references.

IN THE NEXT ISSUE: HEALTH CARE AND MINORITIES

ED185193
UD020455

INTRODUCTION: A LAWYER LOOKS AT SOCIAL SCIENCE IN THE COURTS

> "The Civil War Amendments did not turn matters that are inherently incommensurable into mere matters of arithmetic."
>
> Justice Felix Frankfurter, in Cassell v. Texas (339 U.S. 283, 291 (1950)).

> "In the problem of racial discrimination, statistics often tell much and courts listen."
>
> Judge John R. Brown, in Alabama v. U. S. (304 F. 2d 583, 586 (5th Cir., (1962)).

In this issue of the CLEARINGHOUSE, Jim Loewen describes, concisely and in understandable terms, the uses of various types of social science evidence and statistical analysis in contemporary race and sex discrimination cases. The discussion is designed to familiarize civil rights lawyers with an array of recognized techniques available to marshall and interpret evidence in ways that will assist a court in discerning the wrong alleged. The issue may also help to demystify the legal process for social scientists, providing them with a better understanding of how their research can serve the needs of judicial decision-making.

These concrete and current illustrations of social science in action should be viewed in the perspective of continuing controversy over the role of social science in court proceedings. That role, of course, is not of recent derivation. Nineteenth and early twentieth century decisions of the Supreme Court are replete with unacknowledged social science judgments, as when Justice Brown, speaking for the Court in Plessy v. Ferguson (163 U.S. 537 (1896)), upheld state segregation statutes with the observation that if such laws stamped the colored race with a badge of inferiority,

> ...it is not by reason of anything found in the act, but solely because the colored race chooses to put that construction on it.

Similarly, in cases such as Lochner v. New York (198 U.S. 45 (1905)), the Court struck down protective labor laws with statements that "it is manifest" that limitations of the hours of labor cannot be premised upon the need to protect the health of workers, prompting Justice Holmes to remark in dissent that "the 14th Amendment does not enact Mr. Herbert Spencer's Social Statics." Decisions like Lochner led reformer-advocate Louis D. Brandeis to seek to marshall and systematize social science knowledge in briefs demonstrating the reasonable bases on which state legislatures could have relied in enacting protective labor laws. And in Muller v. Oregon (208 U.S. 412 (1908)), the Supreme Court upheld Oregon's law barring the employment of women in factories for more than ten hours a day, distinguishing Lochner on grounds that the Brandeis brief had persuaded the justices that there was ample justification for the

> widespread belief that a woman's physical structure, and the functions she performs in consequence thereof, justify special legislation restricting or qualifying the conditions under which she should be permitted to work. (Id. at 420)

The use of social science in the courts entered a new more controversial phase with the

Supreme Court's decision in Brown v. Board of Education (347 U.S. 483 (1954)). The Court supported its conclusion that state laws calling for racial segregation of public schools had a detrimental effect on black children with the statement that "whatever may have been the extent of psychological knowledge at the time of Plessy v. Ferguson, this finding is amply supported by modern authority;" the Court added a citation, the famous footnote 11, to the writings of Kenneth B. Clark, Gunnar Myrdal, and other prominent psychologists and sociologists. Some of the furor over Brown may have been stimulated by the differing circumstances in which it used social science evidence. In Muller and other cases, Brandeis briefs were employed to sustain the exercise of authority by state legislatures against the Court's tendency to impose its own economic and sociological theories to strike down such laws. In contrast, in Brown and increasingly since, social science has been marshalled in support of efforts to challenge state laws, calling upon the Court to exercise authority on behalf of unpopular minorities. This has opened the way to charges that courts were engaging in judicial lawmaking based upon nothing more than social science conclusions that were controverted and ephemeral.

Most civil rights lawyers (this one included) contend, however, that the Brown decision did not rest on the social science authority cited in footnote 11. Rather, we believe that the basis of Brown was the fundamental moral judgment made in the Fourteenth Amendment -- that black people had a right "to exemption from unfriendly legislation" (Strauder v. West Virginia, 100 W.S. 303, 307-308 (1880)) directed against them because of their race, that racial classifications in this society can be upheld only if they meet the heaviest burden of justification. Social science authority was cited by the Court only to add weight to the decision -- to persuade people that it was important for the Court to take this difficult step because the lives of children were being damaged by segregation laws. Ironically, it was this effort to foster popular acceptance of the decision that led to the greatest controversy.

If civil rights lawyers argue that sociology was not in fact or properly the basis of the Brown decision, how can we contend that statistical analyses and other social science evidence should play a central role in determining whether voting procedures, employment tests, or similar practices violate civil rights laws?

A simple but vital distinction must be made. In Brown, the Court was dealing with state laws imposed by a dominant white majority on a politically powerless racial minority. The issue was whether such laws, explicitly classifying people by their race could be squared with the fundamental values embodied in the Constitution. Most of us would agree with Edmund Cahn that such a crucial decision should no more be made to depend on shifting social science judgments as to whether the classification was harmful to those affected by it, than persons denied the right to vote or to the aid of a lawyer in a criminal proceeding should be put to proof that such deprivations caused them measurable injury. In the cases discussed in this CLEARINGHOUSE, the issue is whether policies or practices that are not on their face racial in fact are governed by racial considerations or, alternatively, whether nonracial justifications offered for practices that have racial effects are well grounded and persuasive. In such situations where facts are ambiguous, their implications unclear, or relevant information is lacking, experts can be of real assistance to courts in interpreting the evidence and reaching conclusions. Statistical disparities that appear to the untrained observer to be insignificant, upon the application of recognized techniques of analysis may turn out to be meaningful (or vice versa). Employment tests that appear to be valid predictors of job performance may in fact have major deficiencies.

Social science can be useful not only in discerning the role that race has played in challenged practices but in aiding the search for legal remedy. In race discrimination and other cases involving public questions, the role of courts is to apply "equitable principles," and equity, the Supreme Court has said, is "characterized by a practical

flexibility in shaping its remedies and by a facility for recognizing public and private needs." (Brown v. Board of Education (Brown II), 349 U.S. 294 (1955)) In a school desegregation case, for example, numerous questions may arise at the remedy stage on which the expertise of social scientists and educators can be useful. Of a number of proposed plans, which is likely to prove the most stable and enduring? Which reassignment plan is best calculated to overcome educational disadvantage? What other compensatory or remedial measures are needed to overcome the effects of past segregation and establish a climate for equitable treatment in the newly desegregated school? On each of these matters, judicial discretion may be limited by legal principles, such as that community opposition cannot impair the constitutional rights of minority students, but within those bounds most judges are eager to obtain the guidance of experts and experienced observers.

There are risks, of course. The lure of near certainty that statistical methodology offers may lead to judgments that are neat, precise, and wrong. Some social science conclusions, such as those concerning women in the Brandeis brief quoted in Muller, may in retrospect appear very perishable. Yet except where social science is used by judges to evade the need for adhering to fundamental constitutional values, even perishable analyses may be preferable to uninformed judicial theory based upon mere speculation.

Indeed, the principal constraints on the use of social science evidence in the courts may be those that exist within the profession itself. At most universities, there is far more academic payoff in publishing in a prestigious research journal then in preparing even careful and well-crafted testimony for a lawsuit. Some social scientists may be loath to face the cross examination of uninformed but overbearing lawyers, to be bound by narrow rules of legal evidence, to draw conclusions on matters that need study for another decade. Yet "best judgments" are often desperately needed in legal policy-making, especially when definitive conclusions are not yet available. And, for those who still hold such reservations, an admonition of Holmes may prove useful:

> "It is required of a man [or woman] that he should share in the passion and action of his time on peril of being judged not to have lived."

Volume 1 #2 of Jobs Watch is now available from the Center for National Policy Review, Suite 507, 1725 Eye Street NW, Washington, DC 20006. It is free of charge. This issue focuses on sex discrimination in employment.

HOW SOCIAL SCIENTISTS AND LAWYERS CAN WORK TOGETHER

Importance of the expert.

Until recently, the key elements required to obtain legal redress of a wrong arising from racial, sexual, or other discrimination were three: a courageous plaintiff, an informed lawyer, and the development of the law. The factual basis of the case was usually clear. In Mississippi, for instance, not one black was a highway patrolman in 1968. Discrimination was obvious.

Today, North and South, the factual situation is subtler. No longer does a large company have no women as higher managers; instead, perhaps 17% of its management are women, and that proportion may or may not indicate discrimination -- other facts must be known to place it into context. So a fourth element is often required today in law suits to realize civil rights: participation of a social scientist (or several) to make sense of the subtleties in the facts and to counter the inevitable arguments from those in charge that no discrimination was involved.

Communication across the boundaries of lawyer, social scientist, community activist, and plaintiff can be difficult, however, for disciplinary boundaries are involved. Sociologists speak a statistical language that can only occasionally be comprehended by other sociologists; rarely do economists write prose at all; and both groups are put off by legal jargon. Furthermore, the enterprise of studying a problem in order to develop court testimony or exhibits is foreign to most social scientists, upon whom it places new demands for quick response and near-certainty.

On the other hand, many social scientists would like to make more of an impact upon the world than their ivory-towered publications allow, and they generally would wish that impact to be in the direction of broadening opportunities, rather than increasing the resources available to the privileged. More lawyers might seek and use social science testimony if they knew its uses and how to solicit it. And many community activists and institutional leaders could use social science evidence to bolster their programs if they knew where to locate it.

Statistical evidence and tests are usually an important part of social science testimony. Statistics is a branch of mathematics; it is therefore in most simple applications an exact study; accordingly, two statisticians will usually not disagree on calculations based on the same formulas and data. There are two wide areas for disagreement, however. First, questions arise as to which statistical test to use and whether its assumptions have been met. Second, what variables need to be considered and controlled for before causation can be inferred? Different answers to these questions can lead to different conclusions by different experts, sometimes frustrating lawyers and the court.

Social science experts should be equipped to deal with both of these areas of potential disagreement. They should know when the assumptions upon which a test is based can be violated without damage to the test's power or utility, and they should know what additional variables a social analyst should control for in order to establish a relationship. When they can address these areas competently, they are usually more valuable than "straight" statisticians, for social scientists are conversant with statistical techniques as well as being knowledgeable about the substantive area in which they've been asked to testify, while statisticians are often amateurs in, say, political sociology. But since the statistical underpinnings of the expert's testimony are "exact," s/he can develop tes-

timony that is more persuasive than merely "his opinion," expert though it may be.

These articles will provide some instances of those techniques. They constitute rather easy-to-do researches that have proven courtroom utility. In the presentation of testimony based upon these techniques, it is important to build in the court's mind the (correct) impression that statistics is a science* and that the computed probabilities are the results of exact mathematical calculations, not opinion. It is appropriate, for instance, for the lawyer to ask questions that demonstrate that a given result (say, the exclusion of young adults from juries) could not likely have happened due to chance, that some cause must have been involved, and then to ask, "Is this your opinion, Doctor, or would any competent social scientist or statistician have come up with the same conclusion?" The witness can then reply that virtually anyone would have used the same basic statistical test s/he did and would have come out with identical calculations.

The terms "calculation" and "result" are recommended instead of "estimate". "Estimate" has wide currency in statistics and social science; we use the mean (average) income of a sample of undergraduates, for instance, to form an "estimate" of the mean family income of an entire student body. But in common parlance and in courtrooms unused to expert witnesses, estimate has a meaning that verges upon "guess". To use the term in the courtroom may invite misinterpretation, therefore, since a statistical estimate is not a guess but a precise calculation that eventuates in a range or interval within which the true population income (to continue our example) is almost certain to lie. It is important, then, for the expert to state, "I computed the band of incomes within which the mean family income for the entire student body is almost certain to fall; that band is $29,600 to $34,500." Later, then, the degree of confidence can be added, after that concept has been explained to the court.

Outline of testimony

Testimony or deposition of an expert social scientist will usually follow this outline:

qualifications;
explanation of the nature of the data;
nature of the statistical test(s) performed by the expert; discussion as to why they are the appropriate tests;
presentation of the data and conclusions;
discussion, even speculation, involving interpretations by the expert.

Each of these will now be discussed at greater length.

Qualifications. Move from the general to the particular. Begin with education, highlighting coursework in research methods, statistics, and relevant substantive areas. Move through teaching/research experience, highlighting areas using the statistical techniques or touching the substantive questions relevant to this research. Discuss relevant consultation, publication, and other activities. By the conclusion of this section, the court should perceive the witness as a well-qualified social scientist who has particular familiarity with the literature and research methods related to the topic at hand. A thoughtful, even artful, up-to-date vita buttresses this section.

---

*"Science" is here used imprecisely, since mathematics is not a science (just as grammar is not a form of literature).

Explanation of the nature of the data. Prior witnesses or discovery may have generated them, in which case all that is needed is a reference to the appropriate exhibits and testimony. If the expert gathered the data personally, she (e.g.) should describe in some detail how this was done. If others gathered the data for the expert -- students, perhaps, or the plaintiff(s), or a secretary of the attorney -- it is important that this work has been done under the supervision and at the behest of the expert witness. Therefore the social scientist should talk with the gatherers before they begin, telling them what to watch for, what to collect, etc. Then they become agents under her supervision; their interviewing techniques have been refined by her; they benefit from her expertise. Otherwise the data-gathering process may lose credibility and data may even be inadmissable owing to hearsay. It is strongly recommended that the expert participate, even if only minimally, in every step of the data collection -- that she do some interviews, for instance.

For ethical reasons, it is wrong to submit to the court (or opposing counsel) completed interview forms or questionnaires, unless releases were obtained from all respondents. allowing use of their names and opinions in court. However, names and addresses can be separated from the forms and the forms submitted "naked"; then the other side can check the work without compromising confidentiality. Be sure the names and addresses cannot be subpoenaed; social scientists have gone to jail for refusing to turn over records they did possess, in order to maintain research ethics. Also be sure that persons cannot be identified from their replies, even if names have been deleted. Initial consultation between lawyer and expert can establish procedures so that the social scientist's concern for "informed consent" is met while the attorney violates no rule regarding destroying evidence.

Care must be taken to ensure the admissability of these crucial exhibits containing the raw data. Do not submit a paragraph based on Census data without enclosing a Xeroxed copy of the Census pages on which it is based. Do not submit tables of voting results without enclosing certified election returns, if available. In short, make sure to submit the best available evidence, to ensure the least effective challenge to it.

Nature of the statistical tests performed by the expert and discussion of their appropriateness. To display research designs and data analyses clearly and understandably to the court, a lay audience, is an art. Since undergraduate students and the general public also constitute important lay audiences to whom social scientists must communicate, it is an art well worth cultivating. The expert will want to be economical. Methodological prowess can be demonstrated later in a journal article; tailor testimony to the requirements of the case and the legal situation. The expert should do an extensive and sophisticated analysis, if appropriate, to ensure that s/he is testifying honestly and to protect against unturned stones that might be flipped over by the other side in cross examination. But do not allow "pressure" from having done the analysis to force "publication" in the form of complete presentation to the court of the results. Although the expert may have controlled for thirteen variables at once and developed a path analysis unprecedented in elegance, what is presented to the court may need to be a single 3 x 2 table.

For each statistical analysis to be presented, the expert should prepare a simple one-page exhibit showing the formula, explaining how it works, and citing references in the literature. This can help the lawyer write the brief even if the exhibit never gets presented to the court.

Presentation of the data and conclusions. Data presentation is often the neglected third of methods/statistics courses (which include data collection, data analysis, and data presentation, but often over-concentrate on analysis). Perhaps that is why even experienced social scientists sometimes make elementary errors in data presentation. Specific

errors to avoid and suggestions as to how to proceed are found in the next article. Here let me just suggest: make the presentation clear, simple, and effective. A complex table should be broken down into two tables or two bargraphs or maps. Be sure to include sentences of conclusion or interpretation at the bottom of each figure or table, because the court may later view the exhibit without connecting it to the expert's "prose" (testimony). Indeed, the expert's testimony may be excluded by the trial court, in which case it will have to get in the record for appeal via an "offer of proof," and the expert's conclusions can be more effectively stated on the exhibit than by the attorney speaking extemporaneously about statistics in court.

Discussion. This section needs to be carefully reviewed so that witness and attorney know what to expect. At this point, the data and conclusions have been presented; what is needed is to "make the facts sing," as an attorney once explained to me. Carefully controlled speculation, based on logical and accepted social science postulates, is called for. If the expert has determined that the white elected officials in a county have been elected by a white bloc vote, for instance, while blacks have strongly voted for someone else, it also follows that those officials will likely be more responsive to the needs and interests of the white population than to their black constituents. After all, it is an accepted principle of political science that those who elect someone have more "pull" with him or her afterwards than those who did not. Citations are legion. "Speculation" to that point is hardly speculation but is based on the literature.

In order to avoid the charge of "speculation," the attorney and the witness must carefully restrain the scope of their questions and answers. The lawyer cannot put words into the mouth of the social scientist! And an expert who is expert on everything is expert on nothing. There need to be areas in which the expert does not claim expertise. S/he cannot claim to be equally expert in all areas of social science as they touch the questions involved in the lawsuit, without destroying credibility as a true expert in one of those areas. The lawyer and expert should discuss this point carefully, particularly regarding cross-examination questions. No statement should be risked which the expert does not know to be true. Useful is the phrase, "I have no direct knowledge of that from my own studies, although I do have an opinion," or, weaker, "This is outside the field of my particular expertise, but as an informed social scientist I can speculate as to what most of my colleagues in that area would conclude," followed by a pause. If the judge then asks the expert to go ahead, any objection from the other side has been headed off; if the judge calls a halt, then the judge is in the position of appreciating the honesty of the witness and implying that the rest of his or her testimony has been expert. In sum, the phrase "I don't know" can be a signal of strength, not of weakness.

Tips on Preparation.

Lawyers and social scientists are professionals who face rather different situations and have different needs. Both the lawyer and the scientist have an interest that the expert testimony flow smoothly and be convincing, but ego or other problems can interfere with proper preparation. These problems can be mitigated if each party is aware of the needs of the other.

Most basic is the need each party has to be heard and respected by the other. Lawyers in particular can sometimes inadvertently convey to experts that they are "hired hands," getting paid for their work and working for the lawyer, hence "beneath" the lawyer. While the lawyer does have the responsibility of directing the case, this attitude will not work. Social scientists do not perceive themselves as employees but as consultants, a relationship of equality. Moreover, competent social scientists see their work as partly directed by the data, not the lawyer, so that if they feel that a factual situation deserves complex analysis, they will want to analyze the matter complexly, even if the results are counterpro-

ductive. Social science experts must be able to defend their research before other social scientists, whether on the other side of the courtroom or later, in the form of a tenure review committee or editorial advisory board. The lawyer must have the patience to deal with this characteristic of the expert, who is a professional and is concerned about maintaining professional standards. Indeed, the lawyer should take time to seek out the expert's advice on the best data to obtain, the most powerful ways to do the analysis, and the best ways to present the results.

At the same time, when it comes to the exhibits or testimony to be presented to the court, the expert needs to defer to the lawyer. The expert should remain involved intellectually, suggesting every strategem, pointing out every potential weakness, and raising every objection that comes to mind. But the lawyer is expert on trial procedures and strategy. Although the expert must tell the truth and must not distort the factual reality by errors or omissions, within that precept, the expert should listen to the lawyer's conclusions as to what has to be included, what should be deleted. Be true to the data but be responsive to the needs of the case, and these latter are the lawyer's determination.

At the beginning of the relationship, the lawyer should be clear as to the contractual responsibilities of both parties (and the rate and method of payment). A letter of agreement is called for, for the sake of clarity and because the expert is otherwise going out on a limb in doing the work without guarantee of recompense, even of expenses. Also, the expert may need such a letter to demonstrate that the consultation and testimony did take place. Two-way consultation is then required to make sure that the data that the lawyer will provide the scientist (or that the scientist will generate) will be the best possible data, in final form, gathered ethically. Nothing is more discouraging to the scientist than to complete an elegant computer run, only to learn that some of the figures were preliminary. Because final figures must be used in any analysis to be presented to the court, the lawyer should make sure the expert knows this and delays analysis until they are available.

The lawyer needs to understand the statistical analysis and testimony. The expert must be able to make this analysis clear to the lawyer as a trial run towards making it clear to the judge. The lawyer, meanwhile, must take time to prepare this testimony, so that s/he does not ask statistically inane questions. Generally, if a question on direct examination requires the expert to do any computation on the stand, that signifies inadequate preparation on the attorney's part.

At any break in testimony, it is important for expert and lawyer to provide positive feedback to each other, even if it must be forced. The attorney should be particularly sure to compliment the witness, who will be anxious to know "how am I doing?" Suggestions can then be made in a context of mutual support; "coaching" the witness during breaks must be avoided.

The expert should be cautioned to be brief on cross-examination, not to volunteer a lot, and not to allow harassment. For instance, if opposing counsel interrupts answers, the witness should know that s/he is as much in charge of the conversation as is the lawyer; therefore it may be appropriate to say deferentially, "Sir, you have interrupted my answer to the previous question," and then continue with what was being said. The witness should also be advised to pause a moment before answering, to gather thoughts and to allow counsel to object to the question if appropriate. And if one is an "absentee" expert, whose expertise is based on statistical analysis rather than on-site research, it is correct to reply to the question, "You haven't even <u>been</u> in Smith County, have you?" with a calm non-defensive "no".

After trial, the lawyer should be sure to send the expert a letter of thanks for the successful testimony, which may later be of use to the social scientist's career. The

lawyer might also encourage the expert to publish in a social science journal something based on the analysis, which might then be useful in future cases, either as an exhibit or to bolster the expert standing of its author; publication is also in the expert's career interest. The lawyer can invite the expert to share a draft for prepublication comment, which will be helpful to the author and can also help ensure that unwittingly counter-productive statements are avoided in the article.

If attorney and expert are both prepared, an expert is sometimes indispensable in civil rights cases. The following articles present specific techniques that have been successfully used in court. The expert and lawyer can use or modify them to deal with a wide range of issues likely to be encountered in civil rights litigation.

THE NATURE OF STATISTICS AND RESEARCH AS USED IN CIVIL RIGHTS LITIGATION

Three Branches of Statistics.

Statistics in the courtroom are not different from statistics in other applications, but this article cannot be an introduction to the entire field. Such introductions are available, and I recommend several (Anderson and Zelditch, 1968; Loether and McTavish, 1974; Zelditch, 1959; cf. Freedman, et al., 1978, and Meyer, 1973). This article will constitute a sort of introduction to those introductions; if I am successful the lay reader should come away de-mystified and with a grasp of the basic statistical/research process, particularly as used in the courtroom to convince people of social facts. The social scientist may glean some ways to present statistics clearly to lay audiences.

For our purposes the field of statistics can be divided into three parts: descriptive statistics, measures of association, and inferential statistics. Descriptive statistics summarize an otherwise unwieldy body of information. If we are examining 950 male employees, for instance, we cannot hold in our minds even so simple a characteristic as their ages without using descriptive statistics. The data must be condensed: an age pyramid, table, mean (average), or median (midpoint) might be appropriate. Only then can a second variable be added, such as a comparison to female employees or employees at a different firm. Descriptive statistics comprise the basic information that will be presented to the court and will usually be the most important single part of the expert testimony.

Measures of association are statistics that tell whether two (or more) different variables are consistently related. The simplest example of such a measure is a two-by-two table:

Table I.

Age and Sex at Hyperbolic Manufacturing, Inc. (Fictitious).

	Male	Female
Above 40	494, 52%	108, 40%
Below 40	456, 48%	163, 60%
	950,100%	271,100%

Equally basic is the scattergram; a simple example is shown in "Analyzing Election Returns for Racial Bloc Voting." Several statistical measures have been devised to describe the strength of the relationship shown in tables or scattergrams.

Table I shows a mild association between sex and age: men are likely to be older. Two questions must then be asked about this relationship: what is its importance, and what is its statistical significance. The "importance" of a relationship is a difficult concept to define, because it depends upon the setting. The difference between 51% and 49% is crucial in politics, for instance, but probably trivial when comparing the proportion of whites and blacks admitted to dental school. "Importance" is answered by the seat of one's pants, by common sense, by social theory, by statistical measures of association in contrast to measures between other variable pairs, and by the expert in the light of the literature or other cases. "Significance" is a statistical concept that means "unlikely to occur owing to chance" and is the question posed (and answered) by inferential statistics. In our above example, if we have determined that 52% vs. 40% seems important, perhaps indicating some prejudice against hiring older women or some process that mitigates against female retention, then we must find out if a 12% difference could likely have occurred by chance. (The article, "T Test for the Difference of Two Proportions," shows how to do this.) If chance could likely have produced the results, we accept the "null hypothesis" -- that nothing has been proven. If chance could not likely have produced the disparity, we conclude that sex (gender) or something tied to sex was responsible.

In order to go to court with a finding (or in order to persuade somebody of it in social science), importance and significance must both be shown. Based on a sample of 420,000 workers, a trivial difference between two groups can be statistically significant. And a vast difference -- say 90% of whites are admitted to veterinarian school, compared to only 14% of black applicants -- can be statistically insignificant if based on a tiny sample.

Control Variables.

If one is to convince others of the soundness of one's conclusions, reasonable alternative explanations must be explored and discounted. Another way of putting this is to look for the effects of control variables. An alternate explanation for our sex-age relationship, for instance, might be that the firm has implemented an affirmative action program during the last five years and hired many women in that time, while its male workforce has slowly aged, augmented by fewer new-hires. This is quite a different explanation than the two anti-women processes suggested earlier, but like them it asserts that "sex (gender) or something tied to sex was responsible." Therefore the observed relationship between sex and age can be accounted for by any of these three explanations or by several others; by controlling for a third variable, some competing explanations can be discounted. At the least, the social scientist and lawyer should be able to identify the major competing explanations that exist in the literature or are likely to be put forward by the "other side," and control variables should be investigated so as to be able to comment on each of these explanations. How to control is shown by developing our recency of hiring example.

We might control for recency of hiring by looking among part of the dataset: the recently hired. Table II shows that the sex difference washes out. Among the recently hired, sex makes little difference to age. Neither does it among the previously hired. The apparent effect of sex on age in Table I was really due to recency of hiring; a third table, with hiring across the top and age on the side, males and females combined, would show this hiring effect. With additional theory and investigation and additional controls, it might be possible to resurrect the hypothesis that sexism is rampant at Hyperbolic Mfg., but without more analysis, it looks as if the affirmative action explanation does hold.

Table II. Age and Sex, by Recency of Hiring.

	Recently Hired (within 5 years)		Previously Hired (5 or more years)	
Age:	Male	Female	Male	Female
Above 40	82, 42%	80, 38%	412, 55%	28, 53%
Below 40	118, 58%	131, 62%	338, 45%	32, 47%
	200,100%	211,100%	750,100%	60,100%

Presenting an argument in social science (or making a factual case in court) typically involves the four steps just described:

--after the data are gathered and analyzed, they are summarized effectively using descriptive statistics;

--relationships between two variables (or comparisons between two groups, which amount to a form of relationship between two variables) are presented using statistics of association, including tables;

--important relationships are tested for statistical significance;

--important and significant relationships are subjected to further controls to see if they hold up.

A Mix of Public Methods.

The keenest statistical analysis cannot establish anything if the raw data being analyzed were obtained through indefensible research practices. "Garbage in, garbage out," is a maxim of computer programmers. Good research methods, methods that can convince another social scientist or a judge, are public and mixed.

Methods must be public. That is, the means by which the data were first generated must be able to stand scrutiny for fairness. Any sample must be random (or nonrandom for a reason). Any survey questions must be unbiased. And the whole proof, from research design to final product, must be understandable to the lay public.

An important conclusion should not rest on a simple finding or a single method. Sometimes more than one study by more than one expert may be required. To support a change of venue motion on behalf of a black charged with murder, for instance, all of the following might be appropriate:

--a survey of white attitudes toward white and black criminals, asking sophisticated questions of a random sample of white registered voters.

--a comparison of arrest to conviction-and-sentencing statistics in the county by race

over the past decade.

--a study of the proportion black on recent juries and venires compared to the voting age population.

--content analysis of the treatment of black/white murder compared to white/white murder in the local newspaper.

Some social scientists and users of social science think all too quickly of surveys -- questionnaires or interviews -- when something must be established. Surveys do "create" data, it is true, but that data typically is not about topic X, it is about how people think about topic X. Thus the survey of householders reported in "Selecting a Defensible Sample and Using the T Test for the Difference of Two Means in Municipal Services Discrimination Cases" does not reveal much about city services, but about images and evaluations of those services.

Often data need not be created at all; often it already exists. City planning departments, regional economic development boards, school systems, marketing firms, and other bureaucracies maintain files of potentially useful material (cf. Hauser, 1964; Miller, 1977). Other data are just lying around, waiting to be picked up. If a high-school principal is charged with racism, for example, why bother to construct a questionnaire, select a sample, and ask if black students are being excluded from important extra-curricular activities -- merely count old yearbook photos! Such data sources are called "unobtrusive" (Webb, 1966); they usually bother nobody, so they are ethically superior to surveys. And since their information antedates this use of it, it is implausible to suggest it was biased from its inception.

Creative use of existing data sources and unobtrusive methods help ensure that the data are unbiased and complementary. Careful data analysis assures that defensible conclusions have been drawn from the data. Now these conclusions and data must be presented clearly and effectively to an audience unfamiliar with social science methods and conclusions.

Common Errors in Data Presentation.

When the presentations of social scientists are unclear, it is often because their own training overemphasized data analysis (the controlling and inferential statistics that can often become quite elaborate) at the cost of data presentation. Descriptive statistics are easy; hence they aren't taught. Three errors are particularly common and should be singled out for attention so they can be avoided: non-intuitive mapping, overuse of numerical tables, and mis-percentaging of tables.

A pair of shaded maps offer a rivetting way to show the court the relation between two variables. For instance, students under my direction once shaded the census tract map of Jackson, Mississippi, according to proportion of residential land that was zoned non-residentially -- homes unprotected against commercial or industrial invasion. They shaded black all tracts where this problem was pervasive, left white all tracts where it did not occur, and used greys for intermediate proportions. The map then overlapped almost perfectly with a similarly shaded map of proportion black in the population. The resulting exhibit was more effective than a correlation coefficient between % black and % miszoned could ever have been. No more than five shading categories should be used in any map; they should progress systematically from light to dark. Colors don't reproduce well and should be avoided; grey shadings on plastic film are available from any office supply store. Usually, but not always, the same number of tracts should be shaded in each shade. If there are 40 census tracts, it would make sense to set limits on the variable, proportion black, for instance, such that about ten tracts are shaded "heavily black," ten are "fairly black," ten "fairly white," and ten "heavily white." The next map, showing the variable believed to correlate

with race, can then be shaded with the same four categories, each with about ten tracts, so that if the maps overlap, then the relationship is apparent.

Bar graphs or other "pictorial" representations should often replace numerical tables. The reason demographers draw age/sex population pyramids is because they show graphically the proportions of people in various age categories. If you are trying to show why a 60% black jurisdiction is likely not to have an effective black voting majority, comparative pyramids by race can make almost palpable to the court the dramatic absence of young black adults in Southern rural counties. (See the article on "Racial Bloc Voting" below.) If a photo is occasionally worth a thousand words, a graph can replace a thousand digits.

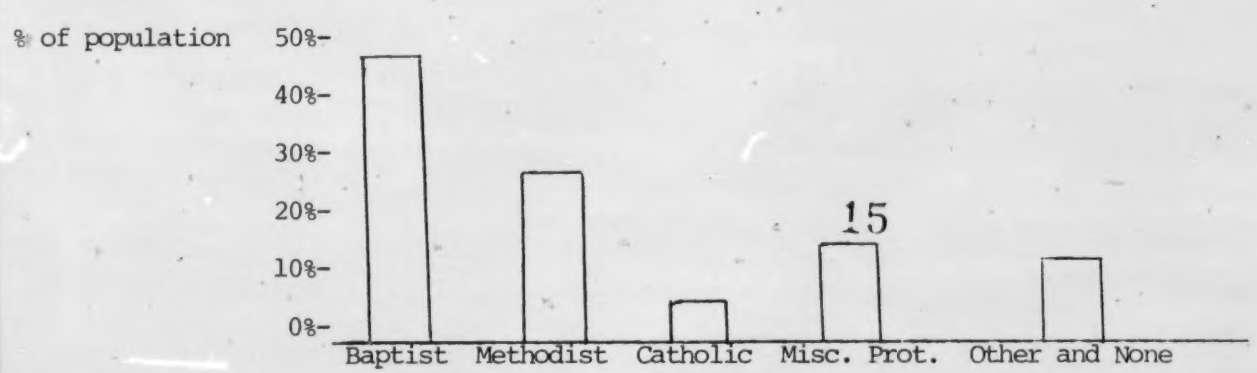
The final error to avoid is mis-percentaged tables. Like bridge bidding rules and like the other admonitions in this article, the following rule is not absolute, but a first principle regarding tables is to put the "independent" variable across the top and percentage the table vertically. Tables I and II are percentaged in that manner, with sex the independent variable which may or may not affect age distribution. Occasionally social scientists percentage tables so that the entire table adds to 100%, rather than each column, or so that the table is laid out horizontally. Sometimes there are valid reasons for so doing; usually there aren't. In Table I, for instance, we wished to know "does sex influence age distribution," which amounts to asking "are women younger than men"? Accordingly, we percentaged to answer, "X% of the women are below 40, compared to Y% of the men." The proportions of women below 40 and above 40 must add up to all women, or 100%; each cell must be divided by the column "marginal", 271.

Parametric and Nonparametric Statistics.

Because statistics are often divided into "parametric" and "non-parametric", we must spend a moment to learn these terms and their implications. Data too come parametrically or nonparametrically. A parametric variable is one that varies over a continuous range with a meaningful zero point. Income, for instance, varies over a continuous range, so that one could earn $10,931 as well as $11,000; $0 is meaningful. Therefore incomes can be added or multiplied. A family with $24,000 has twice the income of one earning $12,000.

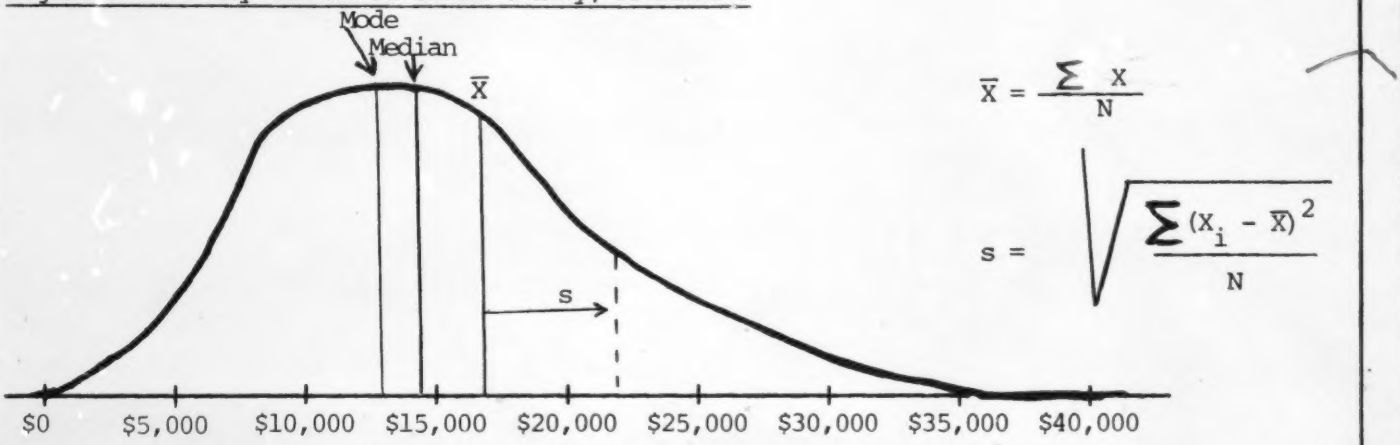
Nonparametric variables vary over a range without a meaningful zero point or a range that is not continuous at all. Religion is an example: if we divide the values of religion as a variable, we might have Protestant, Catholic, Jew, None, and Other. Is a Catholic "more than" a Protestant but "less than" a Jew? Obviously this scale is arbitrary, and we cannot add or multiply its scale divisions. Therefore, we must keep its categories separate, and the only summary we can present would be a bargraph or other representation of the entire frequency distribution, as in Figure 1.

Figure 1. Religious Membership in Shale County, Arkansas.



Parametric distributions, on the other hand, can be represented by "parameters" -- most importantly, the "mean" and "standard deviation". Since we can add all the items together, we can then divide by the number of items to find a "measure of central tendency", the mean, $\overline{X}$. For many distributions, the mean is a good shorthand summary telling where the variable is centered; when the distribution extends out in one direction, the median (the midpoint, found by listing all the cases from smallest to largest and selecting the middle number) would be superior. The second crucial parameter measures the spread or "width" of the distribution and is called the standard deviation or s (see Figure 2).

Figure 2. Family Income in Shale County, Arkansas.



$$\overline{X} = \frac{\sum X}{N}$$

$$s = \sqrt{\frac{\sum (X_i - \overline{X})^2}{N}}$$

Any distribution that looks reasonably like a "bell curve" or normal curve, such as Figure 2, can be described in good detail by its $\overline{X}$ and s. Therefore, a number of statistical inference tests exist which compare two distributions by comparing their parameters -- their means and standard deviations. Other tests exist for nonparametric distributions; among these are chi-square and the sign test. In practice, once the two parameters -- mean and standard deviation -- are mastered, there is little mystery involved in the dichotomy between parametric and nonparametric statistics. Indeed, any variable or relationship that is describable in parametric terms is also describable in nonparametric terms (the reverse is not true). Income, for instance, can be dichotomized into "high" and "low" and treated like Baptist or Catholic. Sometimes nonparametric tests are more powerful than tests of inference based on parametric assumptions and sometimes they aren't; often nonparametric tests are easier to understand; these considerations should dictate the appropriate test for a given set of parametric data. The pages that follow begin with a nonparametric test (sign test), because it can clearly instruct as to the basic ideas of inferential statistics and probability. Then follows the most common parametric significance test, the t test. The articles are written to be read sequentially. Each article introduces a statistical technique in the context of a substantive example, such as discrimination in municipal services. If you have little interest in the substantive topic related in an article, remember that the technique or statistical test has wide application to other topics.

THE SIGN TEST FOR COMPARING HIRING/PROMOTION RECORDS FOR POSSIBLE DISCRIMINATION

Suppose the University of Northern Maryland (fictitious) has the pattern of women faculty members shown in Table I. Obviously women are underrepresented in the higher positions. But the university may claim that women are less qualified, that they have degrees of lesser quality, that they have published less, that their seniority has been broken by pregnancies. In such circumstances, having a set of full professors that are 13% women may represent equal treatment or even affirmative action on behalf of women, the college may argue. How is the court to know?

Table I.	Position	Proportion Women
	Full Professors	13%
	Associate Professors	17%
	Assistant Professors	29%
	Instructors	35%
	Part-time untenured	15%

The Sign Test (Siegel, 1956, 68-75) can be used to test these countering claims. This simple statistical procedure is economical of time and resources, easy for judges or juries to understand, applicable to a wide variety of situations, and constitutes powerful evidence of the presence or absence of discrimination.

Basically it is analogous to a coin flip. For example, two professors, one of each sex, having identical ranks (associate, tenured, for instance) are probably not identical in actual qualifications. One is surely more qualified than the other, perhaps considerably more. Like a single coin flip, either the male or the female professor has to be more qualified, hence underranked, although if the difference is small, the underranking would not amount to a full difference in position and could justifiably be disregarded by the college. If the woman in our example happened to be more qualified, hence was underranked, no one would assume institutional bias, just as no one would assume a coin to be biased if a single toss came up heads.

If nine out of ten flips came up heads, no reasonable person would doubt that the coin was biased. Inferential statistics here parallels common sense, for the likelihood of nine heads in ten flips is a mere .011 or 1.1%. That is, if we engaged in 100 series of ten flips with an unbiased coin and an unbiased flipping procedure, only one of those series would come out with nine or more heads. Far more likely would be five heads or four, or six or even seven.

Statisticians and social scientists use this kind of process, inferential statistics or statistical inference, to rule out "chance" as a likely cause of the result they have observed. Thus if we amassed ten pairs of vita (resumes), each pair consisting of a woman and man of equal rank, we could reasonably expect by chance that in about half of those pairs the woman would in fact have superior qualifications, and in half the man would be superior. If the woman's qualifications were superior in nine or ten out of ten pairs, then

to conclude that this happened by chance would be wrongheaded, for it would occur by chance only 1.1 times in 100.

Through discovery or other means, obtain the vitae or resumes of all employees. Xerox them, delete names and all references to sex, and re-Xerox. These documents can now be compared "blind" by "referees" who are obviously unbiased because they will not know the sex of the individuals whose vitas are comparing. Match each woman with a man of the same current employment rank within the department or area; if more than one man is available, it is important to select your match randomly. (A later article, "Selecting a Defensible Sample" discusses random selection.)

Now give them to the expert, who should be someone familiar with college administration, perhaps a former dean or department chairperson. S/he is then to choose the more outstanding vita of each pair. If there is no significant difference, a tie is reported and that pair is dropped, cutting the N.

There are obvious refinements that can be added. For instance, you could use multiple referees. Two persons could evaluate the entire sample, or someone with credentials in natural science could evaluate persons in that area, someone in the humanities could evaluate those teachers, and a social scientist could evaluate social scientists. You might be in a position to evaluate not just the vitae but the entire promotion/tenure files of each faculty member, including copies of publications, teaching evaluations, and other data.

To return to our small example, suppose that this blind refereeing showed that in nine of ten pairs, the woman's qualifications were superior. Something besides chance was involved, and while we cannot say that outright sex discrimination has been proven, we can say that whatever factor was involved correlates highly with sex. Thus the plaintiff has shown a statistical pattern of discrimination, shifting the burden to the defendent to establish that there is a nondiscriminatory explanation for the pattern. But owing to the fact that the analysis was blind, the expert has already anticipated any such rebuttal.

That last statement bears repetition and clarification. The expert was not privy to the thoughts of the promotion-tenure committee or administration. So s/he cannot tell whether sex was a factor in promotion or if perhaps height was. Women are shorter than men, and if the committee liked height, female shortness may have held women back. But something held them back, something that is not job-related or at any rate does not show up on the vitae or other documents. Educational background, for instance, does show up on the vitae. It might be alleged that women went to inferior colleges or did inferior graduate work, and that's why they were passed over. However, if the expert looking at anonymous vitae did not recognize this inferiority, then it probably does not exist. Not only do you have the expert's reputation as an expert, but you also have the blindness of the procedure to increase credibility; the vitae themselves (with names removed) become an exhibit so that the judge can make the test himself/herself.

Now, even if there is substantial discrimination, there may not be a marked difference in the direction of female overqualification in nine out of ten cases! Therefore a larger sample is appropriate. Table II shows the level of significance to be obtained from different size samples. "Level of significance," the crucial concept of inferential statistics, means the probability that a given outcome (or one even more extreme) could have occurred by chance. In our example, obtaining nine or ten cases of female overqualification ("heads") out of ten matched pairs ("flips") would happen by chance 1.1 times in 100 attempts; the level of significance is .011 or 1.1%. In social science we usually like to have at least "the .05 level of significance," meaning a result that would occur fewer than 5 times in 100 by chance. Our example, nine heads in ten flips, meets this criterion, and almost meets the more rigorous ".01 level of significance" that constitutes another common

Table II. Probability or "Significance Level" for Selected Results of the Sign Test.

Result (# of positive outcomes or "heads")	Trials (# of flips)	Probability of that result or one more extreme, by chance
6	10	.377
9	10	.011
60	100	.029
90	100	less than .00001

benchmark in social science. Table II indicates that a larger sample makes significance easier to obtain. This statistical or mathematical result is again parallel to common sense, for although we might not conclude that a college was unfair if 7 of 10 vitae showed female overqualification, we would (and should) if 70 out of 100 did so. The level of significance, the probability of a result as extreme as observed or more extreme, is calculable from what is called the "binomial expansion," but there is no need to know what that is; statistics books have these tables built in (Siegel, 1956, 250).

It is hoped that many other uses of this simple technique will suggest themselves. Obviously, initial hiring is amenable to the same treatment. It has been claimed that black department heads and principals were placed under less qualified whites after court-ordered faculty desegregation in many public school systems; the Sign Test could easily be used to see if this was the case. In areas outside employment, the applications are equally wide. Allegedly discriminatory tax assessments could be treated by it. To generalize, on any occasion of alleged discrimination, if you can find a reasonable match for each person in the class of alleged victims, and if you cannot quantify the nondiscriminatory grounds on which the decisions were allegedly made, but can only assess them comparatively and imprecisely, then the Sign Test is appropriate.

Always the aspect that is difficult to measure precisely is the aspect to be measured comparatively. Sex and rank are precise: qualifications are not. Housing assessment in the tax roles is precise; value of house (appraisal) is not. It is hard to establish objective standards for comparing the many elements that constitute good qualifications for a faculty member. It is much easier and more defensible to claim that vita A was clearly superior to vita B, and that vita D was better than C, and so forth, particularly with the added safeguard of blindness, so the expert could not have exercised bias even if s/he had wanted to.

THE T TEST FOR THE DIFFERENCE OF TWO PROPORTIONS: ANALYZING JURIES AND VENIRES FOR DISCRIMINATION

Juries are to be composed of peers of the person on trial. This has not been construed to mean that left-handed Irish-American bartenders get to be judged by left-handed Irish-American bartenders but rather that juries should constitute a representative cross-section of the adult members of a community. Such samples should not be biased by race, sex, age, social class, or any other salient social characteristic. If a defendant is to be judged by a jury whose composition is biased against him on one or more of the above dimensions, the bias itself can constitute evidence of unfairness and thus provide grounds for appeal.

A random sample is representative. Indeed, it represents the population from which it was drawn in every regard. Therefore if a reasonably large venire is drawn randomly from the adult population of a community, it should approximate the average height of that population, its range of ideologies, and so on, in addition to its racial, sexual, etc., makeup.

Often a random procedure is allegedly used, but on closer examination it proves to be not random at all. For instance, the same persons may be drawn repeatedly, to constitute at least some members of a jury. The probability of this happening randomly is very small. In at least one Southern county, this device has been used to bias juries against blacks, for although each jury was racially balanced, the same blacks recurred, blacks who could be relied upon to follow the lead of the white foreman. Alternately, a random procedure may be in place, but the underlying population of names from which it selects may be biased, may not be a true listing of the adult population. If a two-year-old voter list is used for jury selection, all adults ages 18-20 are left out, and since young people are less likely to have registered, many persons ages 20-24 will also be omitted. Exclusion of college dormitory addresses also excludes young adults from juries. But in many "victimless" or "lifestyle" crimes, if a jury is composed mainly of older adults, with scant representation of what might be termed youthful points of view toward homosexuality, marijuana, etc., then the defendant faces bias and is not being judged by a jury of his peers, a representative selection of the adults in the community.

To remedy the situation, you need not attempt to learn, through discovery or deposition of hostile witnesses, exactly how jury lists are drawn, even though that information would be helpful. You can simply test the results of the procedure against "chance".

Suppose in Robertshaw County (fictitious, but the Census data are real), whose county seat is a college town, the 1970 adult population contains 8949 people age 18-24, 47% of the adult population, while the 1978 county master jury list (MJL) included 108 young adults (18-24) and 432 older adults. We wish to apply inferential statistics to draw conclusions about the likelihood of a given outcome from specified original conditions. In this example, the basic analytic problem regarding possible jury bias is the selection of a sample from a larger base population. This problem is common to many other situations. For instance, suppose a police force is 32% black, while the total labor force of the community is 59% black. Or there may have been 130 infant deaths in the previous 12 months, of which 52 were minority babies, while minority babies made up only 15% of all births. In each case, a statistician can state that the disparity is too great to have occurred by chance; something else must be involved.

Table I. Voting Age Population (VAP) and Master Jury List (MJL), Robertshaw County.

	VAP	MJL
Young adults (18-24)	8,949; 47%	108; 20%
Older adults (25 and over)	10,126; 53%	432; 80%
	19,075; 100%	540; 100%

The statistical test for analyzing whether a sample proportion is significantly different from that of the population is the t test for the difference of two proportions; its formula is:

$$t = \frac{\text{\% young in VAP} - \text{\% young on MJL}}{\sqrt{\frac{(\text{\% young in VAP})(\text{\% older in VAP})}{\text{total number on MJL}}}}$$

All percentages are to be expressed as decimals. In our example,

$$t = \frac{.47 - .20}{\sqrt{\frac{(.47)\ (.53)}{540}}} = \frac{.27}{.021} = 12.6.$$

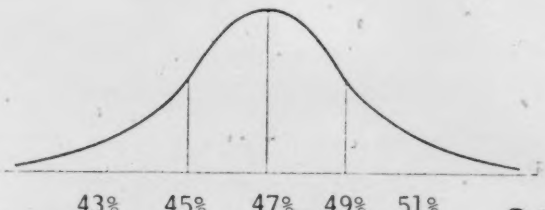
To see what this t value means, consult a t table, in the back of any statistics text; this t value means an MJL containing 20% or fewer young adults would happen by chance far less than one time in 100,000.

Let us take time to understand this formula and example more fully, since it is basic to other related tests, such as the t test for difference between two means. Note that the formula involves a difference divided by a number, the latter expressed in square root form. All t tests share this characteristic. The difference is the population proportion (or mean) minus the observed sample proportion (or the observed sample mean). Subtracting 47% minus 20% leaves 27% or .27, but we do not immediately know if that difference is great or trivial. Common sense (or social theory) indicates that a jury composed of 47% young adults (perhaps 6 of 12) might behave differently from a jury about 20% young (2 or 3 of 12). So the difference seems important. But what about statistical significance? Is the difference meaningful statistically, or could it easily occur by chance?

As with our coinflip examples, the answer depends upon sample size (N). 540 seems to be a large N, so again the statistical conclusion (less than once in 100,000 by chance) parallels common sense. But how did we arrive at it?

A sample will not exactly mirror the population even if taken randomly. If the population is 47% young adult, a random sample of 10 might include 5 young adults, or 4 or 6, perhaps 7. A sample of 100 would include about 47 young adults, and although 48 or 51 or 43 persons might be young, it would be unlikely to find 60 or 38 young adults in a random sample that large. If you took repeated samples, their young adult proportions would come out about 47%; we could create a distribution of those proportions that would look like this:

Figure 1.



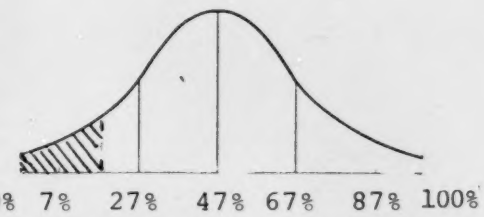
$\bar{X} = 47\%$

$s_{\bar{x}} = 2\%$

This is the famous "bell-shaped curve" or "normal curve" and it indicates, like common sense, that 47% is a likely outcome of a sample, 49% is fairly likely, and 20% (our observed outcome) is very unlikely!

A normal curve has two key characteristics or parameters: mean and standard deviation. The distribution of sample proportions will cluster about the population proportion, 47% in this case. The standard deviation tells how spread out the curve is; Figure 2 shows a curve with mean 47% but standard deviation of 20%. The shaded portion of the graph indicates the likelihood that our outcome of 20% young could have occurred by chance on this broader curve, and we see that that likelihood, though small, is not inconsequential; perhaps 8% of the area is shaded. This curve would result, incidentally, if our sample size were a mere 6.

Figure 2.



$\bar{X}$ = 47%

$s_{\bar{X}}$ = 20%

The standard deviation of a proportion is computed by the denominator of the formula, so our t value actually tells how far out from the mean proportion is our observed proportion, in standard deviations. The t table then tells how likely such a result is, visually equivalent to the shaded area on Figure 2. Because our sample was large, its standard deviation was about 2%, shown in Figure 1, and the probability of obtaining a representation of young adults of 20% or less is vanishingly small, so there is no shaded area on that figure. Any statistician should conclude that the selection was not random but was instead influenced by age or by some characteristic itself associated with age (such as having registered to vote several years earlier). Age bias has been shown. MJL could not have been drawn randomly from the population. Some process or factor that is biased against young adults had to be involved.

Of course, the analysis is not yet complete. We do not know what the biasing factor was, and we may never know, for it may be simply a capricious clerk who throws out all names with university addresses in our college town. But we must check out some prime alternatives. It may be claimed that the names were drawn from the list of registered voters, for instance, and that since students are transient, perhaps maintaining registration and legal residence in another locality, young adults don't register locally. Although that may not excuse the resulting jury bias, it is quite a different process from a capricious clerk. So we obtain access to voter registration cards to find birthdates and calculate ages. We can do this for a random sample if it would be too arduous to skim all the cards; if we do select a sample, we must slightly alter the formula (see Anderson and Zelditch, 1968). Suppose we surveil all cards and learn that of 11,440 total, 4920 or 43% are "young". A disparity remains between the MJL and the roll of registered voters; calculating a new t indicates again that chance could not have accounted for the fact of a 20% young MJL from a 43% young registration list.

Uses for this test are legion. For instance, analyze the content of children's literature readers, to see if the proportion of females in stories approximates that of the population. Compare the percentage of women turned down for credit with that of men. What pro-

portion of units have plumbing problems in black- compared to white-occupied public housing?

Our Robertshaw County example illustrates that some bias had to be involved. The social scientist proceeds by considering how probable the observed outcome would be if selection had been fair and random. If s/he finds that probability extremely low, s/he concludes that an alternative hypothesis -- selection directly influenced by age or by a characteristic itself tied to race -- must be accepted. It is very hard to refute this testimony, properly understood.

SELECTING A DEFENSIBLE SAMPLE AND USING THE T TEST FOR THE DIFFERENCE OF TWO MEANS IN MUNICIPAL SERVICES DISCRIMINATION CASES

Most people who aren't social scientists put too much credence in large samples. Research based on large samples, such as Equality of Educational Opportunity (the "Coleman Report"), is accorded undue praise and importance; when laypeople (and some social scientists) do their own research, they tend to choose too large a sample. Students of sampling know that the famous survey fiascos of the past were based on samples that were too large, not too small. For example, when the Literary Digest predicted in 1936 that FDR would lose, the prediction was based on a sample of more than 50,000 persons; Gallup's prediction of Dewey over Truman, leading to the famous photograph of a grinning Truman holding up the Chicago Tribune front page headline heralding Dewey's victory, was based on many more respondents than his more carefully selected samples of recent years.

How can a sample be too large? If it is not random (or if its nonrandomness is not understood and intended or compensated for by the social scientist), then its size is no guarantee of quality and may even mitigate against quality.

How large should a sample be? That depends upon three factors:

--cost (including such nonmonetary considerations as time, personnel available, and bother to respondents, if a survey is involved).

--plans for data analysis.

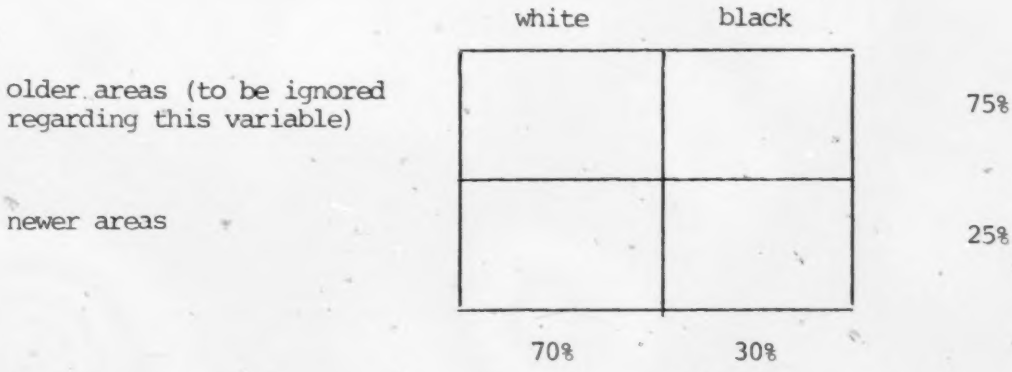
--the need for representativeness and defensibility.

Cost, broadly construed, is a greater problem than sometimes appears. Many research projects have lacked effectiveness because a disproportionate amount of time and energy was spent on data gathering, owing to an overly large sample. The research process includes most of these steps (and sometimes others): consulting with the attorney(s) and client(s), deciding upon a research design, developing a questionnaire (if a survey is the appropriate method), pre-testing it, then selecting a sample, interviewing its members, re-visiting those absent, entering the data into the computer, completing a run, locating errors in the data, correcting them, completing a second run, analyzing the results, interpreting them, creating effective ways to present the data, and consulting again with the attorney(s) and client(s), not only at the end but throughout the process. There are far too many other steps to allow the interviewing, re-visiting, and data-entry steps to dominate the enterprise, which they will do if the sample is large.

As a rule of thumb, a sample size numbering 30 ("N" = 30) provides stable statistics

and allows for percentaging. Suppose, for instance, that a company claims to have no data on the race of persons it has rejected for employment, but has name/address/phone numbers for them. Suppose further that the company is located in a county with a 35% black workforce, while its employees are only 8% black. It would be tedious to learn the racial identities of all 1818 disappointed job-seekers of the past 24 months at the factory, but if in a sample of 30, chosen at random, 10 are black, it indicates that the proportion, 33% black, is not much different from the workforce in the county but is markedly different from the workforce at the plant. Therefore, even so small a sample shows that lack of applications is unlikely to be the problem; some form of rejection after they apply is differentially affecting black applicants.

This need for an N of 30 applies to subsets of the sample if these are to be analyzed separately. For instance, to continue the above example, if we wished to compare qualifications or test scores of white versus black applicants, then we need 30 in each group, or 60, and if blacks make up 33% of the population of applicants, to obtain 30 blacks in a random sample of applicants requires an N of 90. More complex internal comparisons may be required, necessitating still larger samples. For instance, it may be claimed that a city is not providing its newer black subdivisions with street paving equal to that provided white subdivisions. Some older residential areas of the city may have been built for whites but are now inhabited by blacks; it is not claimed that those areas are afflicted by narrow or poorly paved streets. Accordingly, we are now in a four-cell research design:



If the city is 70% white and if one-fourth of its streets are categorizable as "newer," then to have 30 city blocks in the "black newer" cell requires a total sample of 400 blocks. Of course, there are also acceptable ways to oversample a particular category deliberately so that fewer than 400 blocks will nonetheless yield 30 "newer black" blocks.

Gallup and other national pollsters predict the behavior of the American electorate from samples of 1600 people, or .0007% of the nation's population! This small sample is more accurate than earlier larger ones because elements of nonrandomness in earlier samples have been carefully eliminated or reduced. Nonetheless, the public tends to believe that "more is more," and since judges are members of that public, a study of street paving based upon 30 white and 30 black blocks in a city of 300,000 would probably not be convincing, even if statistically sound. The sample neglects the entire X neighborhood, it might be charged. So to be convincing, a sample needs to be representative -- large enough, in this instance, to include blocks from each major part of town. Finally, it needs to be defensible. The defense will partly be statistical, along lines already discussed, but sample size must ultimately be a strategic decision. An instance of sampling in a municipal services investigation I conducted provides some pointers on sample size and on how to take a sample properly.

The task was to study the municipal services of a metropolitan area much larger than

Shaw, Mississippi, site of the landmark municipal services discrimination case of Hawkins v. Shaw (437 F. 2d 1268 [5th Cir. 1972]). The entire town of Shaw (pop. c. 2500) could be surveyed and was, so no sampling was required. But in a city of 175,000 a total survey would be prohibitively expensive.

In order to decide what N would be strategic, working with Dr. Mickey K. Clampit I determined how many blocks of streets there were in the city. This was no small task, requiring verification of the city map by driving perhaps one-fourth of the total street mileage, glimpsing all side streets. "One block" was given an "operational definition" different from a Census block, which is appropriate for population but not for analyzing street characteristics. On the corrected map every block received a unique number; the number of blocks in the city was found to be 6600. A sample must always relate to a population, which it is then supposed to mirror accurately, and our population N was 6600. We decided to take a sample of 660, one in ten, because we felt it could be handled by the student assistants available to us and was "logical" to ten-fingered judges; for purposes of statistical analysis, an N of less than 200 would have sufficed. If I were repeating the project, I would suggest an N of 330, sure to yield 200 residential blocks, for we did find the data-gathering exhausting.

The simplest way to take a sample is randomly. "Random" is a word often bandied about, as in "I took a random sample of undergraduates at the Student Union." Such a sample is not random, of course. "Random" has a specific meaning: a random sample is drawn from the underlying population in such a way that every member of the population has an equal chance of being included. (Not every student has an equal chance of being at the Student Union; commuters, for one, may be undersampled.) A simple random sample is correctly taken by assigning each member of the population a unique number and then using a random digits table, available in the back of most statistics texts.* A random sample is crucial because it can be assumed to resemble the total population in all respects. Therefore if a relationship is unearthed between race and unpaved streets, no one could effectively claim that the sample only included rich white blocks, for instance, because a random sample would include rich and poor white blocks in proportion to the population.

After sampling, we surveyed every included block with two instruments: a form reporting physical characteristics such as street width, type of drainage, and number of street lights, and an interview schedule asked of one household on each sampled block, asking about fire protection, garbage pickup, etc. We had set up these forms to be easily computer analyzed and had pretested them for clarity and brevity.

The data showed bias. For instance, the mean (average) street width in the black community was about 29 feet; in white areas it was 41.

Is a twelve foot difference important? Well, parking can hardly be permitted on both sides of the average black street; the protection to sidewalks and children and the residential quietude that parking lanes provide would be lacking in black neighborhoods. Combined with paving and curbing differences, broad concrete streets compared to narrow tar streets provide a substantial increment in property values for white homeowners.

Is the difference significant? The appropriate statistical test is the t test for the difference of two means:

---

*There are other appropriate ways to sample, including "proportional random," "disproportional random" (deliberate oversampling of a part of the population), and "target" sampling; most of them relate back to simple random sampling for their basic justification (see Sudman, 1976).

$$t = \frac{\bar{X}_w - \bar{X}_b}{s_{(\bar{X}_w - \bar{X}_b)}}$$

$\bar{X}_w$ = the mean (average) white street width.

$\bar{X}_b$ = the mean black street width.

$s_{(\bar{X}_w - \bar{X}_b)}$ = the standard deviation of the difference between the means.

Like other t formulas, we see a difference divided by a standard deviation. (If the reader has not already encountered t tests, s/he should refer at this point to the introductory discussion in "The T Test for the Difference of Two Proportions," above.) The standard deviation of the difference of two means must logically be greater than the standard deviation of a single mean. Each mean would vary (if a different sample were taken), so their difference could vary more. The formula for this standard deviation reflects this:

$$s_{(\bar{X}_w - \bar{X}_b)} = \sqrt{s_{\bar{X}_w}^2 + s_{\bar{X}_b}^2}$$

The standard deviation of the white mean street width, in turn, will be small when the sample N is large and will be small when the standard deviation of the variable itself (street width) is small, and its formula reflects this:

$$s_{\bar{X}_w} = \frac{s_{X_w}}{\sqrt{N_w - 1}} \qquad s_{X_w} = \sqrt{\frac{\sum (X_{w_i} - \bar{X}_w)^2}{N_w}}$$

The easiest way to come to terms with these formulas is by working through a small example; in turn, doing the calculations in the form of a table makes them easier. In practice, of course, the calculation would be done by a computer, but there is no substitute for understanding once how it is done.

Table I shows a small sample of four black and six white streets. All we begin with are the street widths; we calculate the mean; we subtract each width from the mean, recording only the absolute (unsigned) value of the difference. Then we square those differences and add them up. Substituting in our formulas:

$$s_{X_w} = \sqrt{\frac{63}{6}} = 3.24 \qquad s_{X_b} = \sqrt{\frac{942}{4}} = 15.3$$

$$s_{\bar{X}_w} = \frac{3.24}{\sqrt{5}} = 1.45 \qquad s_{\bar{X}_b} = \frac{15.3}{\sqrt{3}} = 8.83$$

$$s_{(\bar{X}_w - \bar{X}_b)} = \sqrt{2.1 + 78.0} = 8.95 \qquad t = \frac{41.1 - 29.0}{8.95} = 1.35$$

Checking a t table we find that t = 1.3 is not significant. In everyday language, to be less than the white mean by 1.3 standard deviations, as the black mean is, could happen by chance fairly often -- about 10% of the time.* Although the black/white difference was substantial, the sample was too small for it to achieve statistical significance. A larger sample would surely be highly significant.

Table I. Black and White Streets in Small Sample, Michigan.

X_{w_i}	$X_{w_i} - \bar{X}_w$	$(X_{w_i} - \bar{X}_w)^2$	X_{b_i}	$X_{b_i} - \bar{X}_b$	$(X_{b_i} - \bar{X}_b)^2$
40	1	1	55	26	676
42	1	1	16	13	169
45	4	16	20	9	81
35	6	36	25	4	16
44	3	9	116	52	942
41	0	0			
247	15	63			

$$\bar{X}_b = \frac{116}{4} = 29.0 \text{ feet.}$$

$$\bar{X}_w = \frac{247}{6} = 41.1 \text{ feet.}$$

When the difference has been found to be significant, then the expert can say with confidence that chance did not account for it. Race or something aligned with it is almost surely the cause of the difference observed. Like the t test for difference of two proportions, this test has many applications. It is usually appropriate wherever the variable (in this case, street width) can be measured parametrically -- that is, on a meaningful numerical scale. This holds true for street width, anything measurable in dollars, time units (such as years of education or days of sick leave), and many other items of social importance.

*This statement is slightly inaccurate for purposes of propadeutic clarity.

USING THE INDEX OF DISSIMILARITY TO DETERMINE THE EXTENT OF SCHOOL OR RESIDENTIAL SEGREGATION

At one time HEW considered a minority child segregated who attended a school in which that minority was in the majority. "Proportion of minority students in majority-minority schools" then became a rule of thumb to compare districts for degree of school segregation. How simplistic! Any school system that was more than 50% black and was totally desegregated, so that each of its schools was identical in racial ratio, would by that rule appear to be completely segregated!*

Social scientists have developed at least half a dozen more sophisticated indices to measure the extent to which a community's schools or neighborhoods are racially segregated. Each has validity for some purpose, but perhaps the most popular and one of the easiest to compute is the "Taeuber Index" (Taeuber and Taeuber, 1965, 195-245; but cf. Becker, 1978). It has the advantage of being unaffected by the overall proportion of minorities in the area. If a city is 10% black but clusters all its black pupils in two 45% black schools, the Taeuber Index will pick this up; some measures would not. Hence the Taeuber Index allows comparison of different districts and of a single district across different time periods.

Its uses in civil rights cases are manifold. Suppose you are suing a school district alleging employment discrimination. It would be useful to show that the district also operates its elementary schools in a segregative manner, to show that a pattern of racism or insensitivity exists and to indicate that prior employment discrimination has had unfortunate effects upon current institutional practices. The Taeber Index offers a quick way to assess the amount of segregation among elementary pupils.

The analyst needs the number of black pupils and white pupils at each school. (A decision must be made regarding others: Asian Americans, Hispanics, Native Americans. One way is to leave them out, doing only blacks and whites, and then do the "third" group against whites, leaving out blacks.) D, the Taeuber Index or Index of Dissimilarity, is then found by the following formula:

$$D = \sum (B_i/B - W_i/W), \text{ for all schools where the term in the parentheses is positive.}$$

This formula is not at all forbidding. B_i means the number of black children at the first school; to divide it by B, the total number of black children in the district, yields the proportion of the district's black kids at the first school. If this fraction is larger than the proportion of the district's white children at that school, shown by the second fraction, $\frac{W_i}{W}$, then the result for the first school will be positive and becomes a component of D. Then on to school #2, and so forth. Table I shows the entire process for a district with five elementary schools. These schools are rather segregated, by visual inspection of their enrollments, and the Taeuber Index does not disappoint, coming out a high 82.7. The Index is not only easy to compute but also has a quickly graspable meaning: it represents the proportion of minority students who would have to be transferred to majority schools in

*Readers will recognize that a heavily black but itself desegregated system surrounded by white systems is still segregated in the larger metropolitan context.

Table I. Sakrete, N. J., Public Schools; Analysis for Racial Segregation.

School	# of Bl. Pupils	# of Wh. Pupils	$\frac{B_i}{B}$	$\frac{W_i}{W}$	$\left(\frac{B_i}{B} - \frac{W_i}{W}\right)$
Adam	0	100	.0	.36	negative
Baker	10	90	.045	.32	negative
Chas.	20	80	.09	.29	negative
Dexter	90	10	.409	.036	.373
Efron	100	0	.454	.0	.454
					$\sum$ = .827

D = .827 or, as it is usually expressed, 82.7 on a scale from 0 to 100.

order to obtain perfect desegregation, an Index of 0. (This is not an efficient means to desegregate and is not proposed; any efficient means would involve two-way transfers, so that majority students were also being moved to previously minority schools; in that event, the proportion of all students who would have to be moved would be much less.)

So we have here clear evidence of school segregation. Taeuber Index results can then be compared to results for other cities or for the region to give meaning to words like "highly" segregated. Most school systems in the Southeast now have Taeuber Indices of between 7 and 25; indices are somewhat higher in other parts of the country, particularly in large cities (Taeuber and Wilson, 1979).

Housing is more segregated than schooling in America, particularly in the Southeast. Taeuber and Taeuber (1965) gives residential segregation indices for more than 200 American cities, with an update available for 1970 Census data (Sorenson, 1975). Cities range from about 60 to 98, with a median of about 88, and are so highly segregated that no city can be much higher than the median! Therefore it would be hard to single out a particular set of housing projects as particularly segregated compared to the nation, but it would be easy to demonstrate that those projects were highly segregated in the abstract or compared to desegregated parts of the city.

While on the subjects of housing and school segregation, a caution: It has been claimed that most remaining school segregation in America merely results from housing segregation in a one-way relationship. However, a more sophisticated analysis indicates that the relationship between school segregation and residential segregation is reciprocal, with segregated schools helping maintain segregated neighborhoods as well as the reverse (Cf. 38 Social Scientists, 1979).

The Index of Dissimilarity can also be used for many other tasks. Like the Lorenz Curve and its Gini Index, it can be used to measure income inequality. Instead of B_i representing the number of black pupils, I_i would represent the amount of income held by the lowest income category and would be divided by all income in the population under study, or I. D would again vary from 0 to 1 and would represent the percentage of all income that would have to change hands for income equality to prevail. For standards of comparison, D

for income is about 40 for the United States. A higher D can be useful for showing that the differences among ranks in an industry or university are important and connote considerable income inequality, or for showing that a Black Belt county is marked by unusual income inequality, indicative of a continuing legacy of discrimination.

THE RUNS TEST FOR DETERMINING IF MINORITIES ARE BEING CLUSTERED WITHIN A DORMITORY OR HOUSING PROJECT

A college that is leery of "inflicting" minority students upon its majority undergraduates may know better than to keep minorities out of its dormitories altogether or segregate them into one dorm, but may nonetheless group or cluster them within certain rooms or sections inside its dorms. To the degree that majority and minority students are aware of this practice, it is stigmatizing toward the minority student, and to the degree that minority students are restricted in their choice of rooms, it is discriminatory toward them.

Similarly, a housing project may confine minority families largely to certain buildings or areas, preserving the facade of an integrated policy while maintaining de facto segregation.

Of course, clustering can occur by chance. There is no reason to assume that every dormitory floor would contain the same number of minority students, or that minorities would in every case be assigned majority roommates. Voluntary choice is also a factor in roommate choice, particularly in post-freshman years. Although there often isn't a great range of selections available when one rises to the top of the public housing list, chance and choice play roles there too.

A statistical test exists to tell whether a given sequence of residences shows unusual clustering or could have occurred by chance. (To rule out voluntary choice would require some on-site investigation.) It is the One-Sample Runs Test.

Suppose a coin were tossed 10 times and the following sequence of heads and tails occurred: H H H H H T T T T T. This is akin to a housing project with minorities in one wing, whites in the other. A "run" is a series of identical occurrences; here only two runs occurred! This would seem to be too few for a fair coin or a fair tossing procedure. Non-chance clustering seems to be involved. And the Runs Test Table from Siegel (1956, 252) indicates that statistical inference again parallels common sense: the chance of getting only two or fewer runs when intermixing five items of each type is less than 5 in 100 (5% significance level).

Siegel (1956, 52-58) instructs how this simple test is performed. I used it to determine if Chinese-American students were clustering at Mississippi State University and the University of Mississippi. (They were, but choice was at least partly involved.) The runs test can be used wherever small-scale segregation is suspected, such as when it is claimed that:

--a teacher assigns most minority students to seats in the back of the room.

--a history text treats women, but only in special "women's boxes," not integrated in with the sweep of American history.

--a newspaper prints stories and photos of black weddings, but only once a week, clustered together on a page by themselves.

All of these examples are based on fact, incidentally; each indicates the stigma that is connoted by segregation; each also indicates a potential application for this easy-to-use technique.

USING CORRELATION, OVERLAPPING PERCENTAGES, REGRESSION, AND CENSUS DATA TO ANALYZE ELECTION RETURNS FOR RACIAL BLOC VOTING

Introduction

Because election rules and districts are typically drawn up by persons in power, and because those persons are usually white, those rules and districts are often biased against minorities. For instance, blacks in Shell County, Arkansas (not the true county or state), cannot elect any member of the county board of supervisors. The county is 66% black (1970 Census), but as will be seen, that does not translate into anything like a voting majority. If there were elections by fairly-drawn districts, blacks would likely elect at least two of the five supervisors, allowing them some representation in county decision-making. In 1970, redistricting was ordered, but the supervisors instead switched to at-large elections. The white electorate, which constitutes an effective voting majority, elects all five.

If whites frequently voted for black candidates, then the lack of a black countywide voting majority would not deter black political candidacies. But if whites bloc vote for white candidates, then in conjunction with data showing the Voting Age Population (VAP) in the black and white populations, a social scientist can prove that blacks are effectively shut out from the political process in the absence of redistricting.

This paper will show how to construct that proof. In so doing, it introduces three major ways to test for association between two variables: correlation, overlapping percentages, and regression. These methods complement each other and demonstrate clearly whether there is an important association between the two variables. In this case, all three methods should be used to tell whether racial bloc-voting exists -- whether there is an association between proportion white in the electorate and proportion of votes for the white candidate(s).

Data Sources.

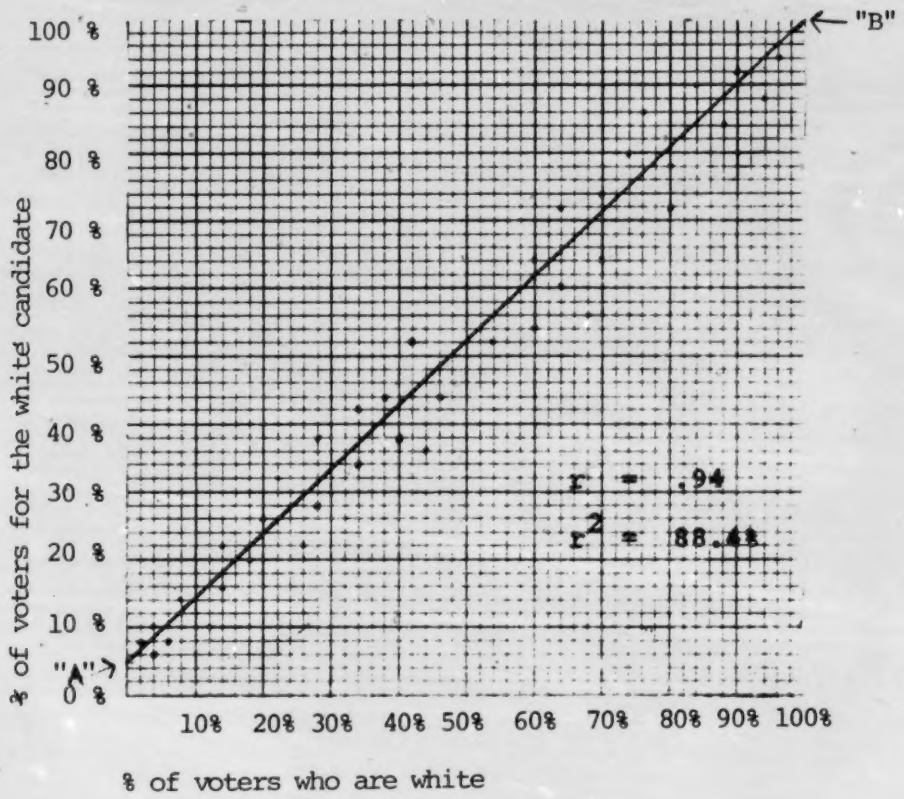
To discover whether two variables are associated, it is necessary to have a measure of the independent variable (the variable that may "cause" or influence the other variable) and the dependent variable (the variable that is "caused"). In the analysis for racial bloc voting, the dependent variable (votes for black and white candidates by precinct) is easy to obtain from newspaper reports or formally filed election returns. The best measure of the independent variable is actual turnout at the polls, by race, by precinct. Next best is the registration roll by race by precinct, as of the time of the election, recently purged. If neither of these is available, then VAP data by race from the United States Census can be used. When using a poor data source for the independent variable, such as unpurged registration lists, it is almost impossible that the problems with the data would be patterned in such a way as to create a spuriously great correlation between race of voter and outcome. Therefore the expert who finds strong evidence of bloc voting even using inferior data

should be able to tell the court that the results would doubtless be yet stronger if better data were available. This is a general principle.

Correlation Analysis.

When confronting two sets of parametric data that may be associated, such as proportion black in student bodies and number of hours of advanced study by faculties, the first analysis that the social scientist should do is a correlation, to see if the two variables are associated. It summarizes the strength of the relationship between variables with a single number, r, and indicates the direction of the relationship with a plus or a minus sign. (A negative indicates an inverse relationship, such as the whiter the precinct, the more votes for the black candidate.) If only a dozen precincts are being analyzed, it is possible to use a hand calculator; otherwise a computer is necessary (see Nie, et al., 1975). Figure 1, a scattergram, is a graphic illustration of this analysis.

Figure 1. Ecological Regression Line.



Each point represents a precinct with the % of voters who are white and the % of votes for the white candidate serving as the coordinates of the point. This particular scattergram shows a strong positive correlation between the % of white voters and % of votes for the white candidate. The line is the "best fit line" or "least squares line" and is a way to summarize the detail and approximate the pattern of the points. The closer the points fall to line, the stronger the correlation.

The coefficient of correlation, r, tells how close the points are to the line; it can vary in size from 0 to 1.0. When r=0, there is no relationship between the two variables; points on the scattergram form a cloud. When r=1.0, the relationship is perfect: the points form a line. In sociology a good r is in the neighborhood of .5 to .7. In voting cases r typically is .8 to .995, astoundingly high. Even more useful than r is r^2, which shows the proportion of variance in votes "explained" by race of voter. The proportion of votes received by the white candidate(s) varies from 6% to 98% in Figure 1; if r^2 is .8, we can attribute 80% of that variation solely to race.

The problem with correlation is that r or r^2 do not tell which group, white or black, is doing the bloc voting and whether they are bloc voting for their own race. The only assertion that can validly be made is that a whiter precinct will produce a greater proportion of votes for the white candidate. To go beyond this, to state "whites are voting white," is to commit the "ecological fallacy" of imputing to individuals a relationship found to exist among groups. Overlapping percentages analysis and ecological regression avoid this fallacy.

Overlapping Percentages Analysis.

Correlation can show us that groups high in one characteristic (e.g., % white) are likely to be high in another (e.g., votes for white candidates). Overlapping percentages analysis allows us to determine whether the individual cases within those groups that have the given characteristic (e.g., whiteness) manifest the expected characteristic in their dependent variable (e.g., voting white). Overlapping percentages allow us to determine how whites and blacks actually voted in overwhelmingly white or black precincts using simple arithmetic. In a heavily white district we start with the assumption that all blacks who voted did so for white candidates (maximum racial crossover). We can then compute the minimum amount of white bloc voting that must have occurred. For example, in a 95% white precinct with one hundred voters the white candidate receives 93 votes and the black candidate 7 votes. Assume that all five blacks voted for the white candidate and subtract that from the 93 votes for the white candidate. That leaves 88 votes that had to come from white voters and means that 92.6% of the whites (88/95) had to have voted white. Even assuming maximum possible racial crossover, the proportion of whites voting white is still very high.

The problems with this method are that it gives us only the minimum amount of racial bloc voting, it is useful for only one precinct at a time, and it can be calculated only for overwhelmingly (more than 90%) uniracial precincts. Its utility goes beyond voting issues, however. For instance, if you know the proportion of applicants from each college winning fellowships for graduate study in a statewide competition, you can use overlapping percentages to see if females or males or blacks or whites were particularly advantaged (so long as overwhelmingly unisexual or uniracial institutions were in the sample). And because it is so simple, it is effective in the courtroom.

Ecological Regression.

Ecological regression provides a better estimate of racial bloc voting or other potential associations between two variables. It can be based on the entire dataset and can be computed without an overwhelmingly uniracial precinct. It works with the "best fit line" of

Figure 1, which approximates the relationship between the race of the voter and the outcome of the election. This line is projected to the point where there are 0% white voters, which is where the line crosses the "y" axis. Point "A" is the point at which there are 0% white voters. The % of votes for the white candidate at this point, 5%, must have come from black voters, since there are no white voters to produce them. By locating the point where there are 100% white voters, "B", we can find the percentage of whites who voted for the white candidate. This figure, about 100%, shows that whites are bloc voting white.

It is also possible to put confidence intervals around the percentages that we obtain from ecological regression. This means that we can draw a band on either side of the regression line and assert that we are 99% sure that the actual line falls within that band. We can also perform a significance test to show how likely it is that a certain association was produced by chance. In conjunction with correlation and overlapping percentages, the result is a powerful statement of the effect of one variable upon the other. Additional refinements, not usually needed in voting analyses, include multiple correlation/regression (looking at the effect of two or more variables at once on a dependent variable) and partial correlation/regression (looking at a variable's effect on another while holding a third variable constant).

Use of Census Data.

Our initial task in voting cases was to demonstrate the association -- to show that racial bloc voting exists -- because if whites are voting for blacks and blacks for whites, then maneuvers of white officials that dilute the black vote and thus decrease the likelihood of black victories will be beside the point, since black candidates can be elected by white votes. The next step is to show why white elected officials are not likely to represent the needs of the black community and then to explain why the black population majority (66% in our example), even given white bloc voting, cannot simply elect its own candidates countywide.

More generally, our initial task in doing correlation/regression analysis is to establish an association between two variables; perhaps some alternative explanations can be discarded through partialling and controlling with third variables. At this point, socioeconomic data on the two or more groups can help explain why the association between the

Table 1. Socioeconomic Status by Race, Shell County (1980 Census).

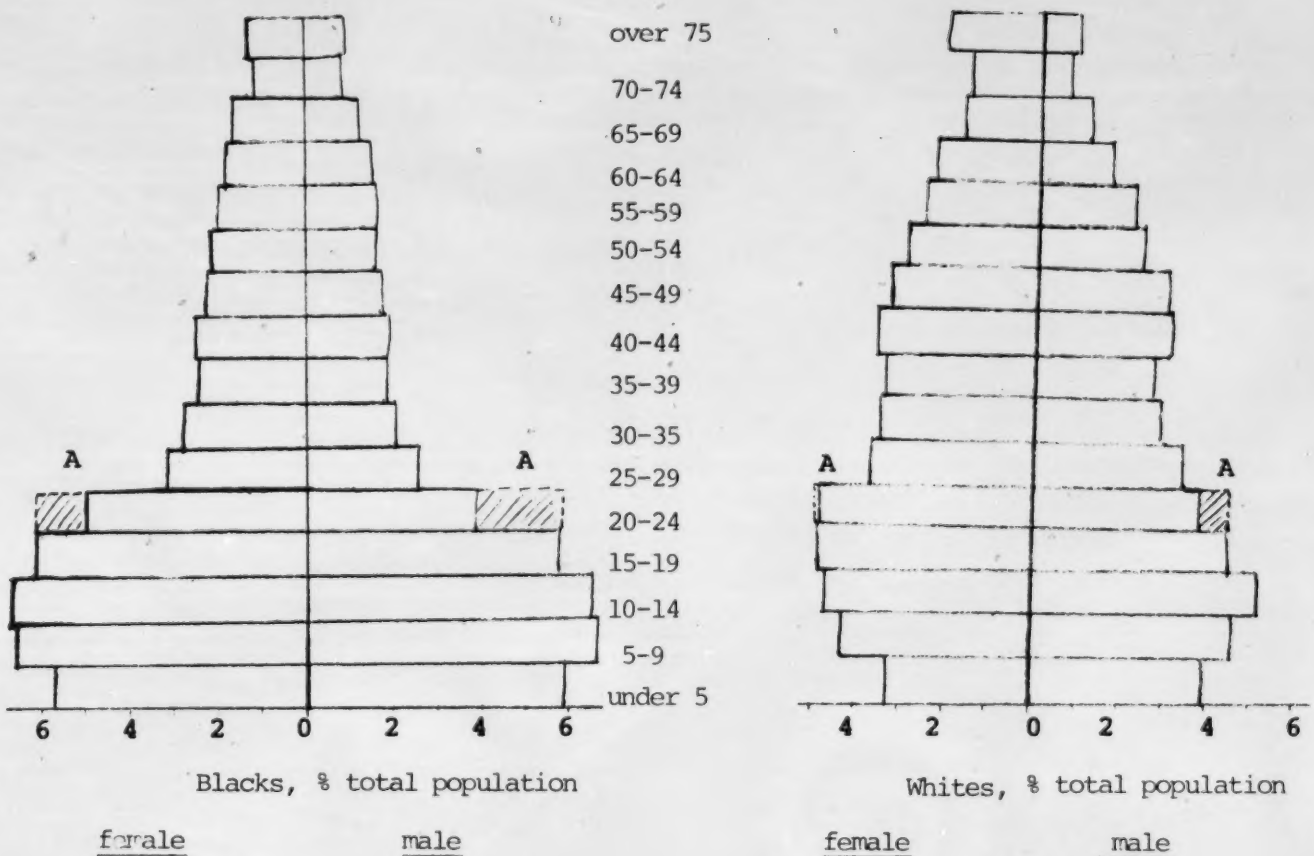
	WHITES	BLACKS
Education		
Median Education	10.4 yrs.	7.0 yrs.
Illiteracy & semi-literacy	4.1 %	27.3 %
Occupation		
White Collar	60.2 %	18.0 %
Blue Collar	39.8 %	82.0 %
Unemployed	2.9 %	7.2 %
Income		
Median Income	$9782	$3794
% below the poverty line	7.3 %	49.0 %
Housing		
% with more than 1 person/room	6.6 %	26.6 %
% families renting	24.8 %	53.3 %

variables exists and can show how that association relates to the overall situation of the disadvantaged group. Census data are the most readily available resource and include information on a wide range of potentially relevant topics, from plumbing fixtures to infant mortality. This section offers an example of its use.

In Southern communities, the median education for whites is typically much higher than for blacks. Most whites have at least some high school education, but most blacks do not. Whites are more likely to be employed in white collar jobs; this, in turn, affects their income level, which is more than twice as high for whites. The physical circumstances of the black community are also inferior. Often most whites are homeowners while most blacks rent. As renters blacks are unable to get their municipal governments to improve services because they lack the clout of direct taxpayers. These differences mean that blacks are likely to have different political goals. The black community is more likely to support public housing and rent subsidies, whereas fewer whites would be eligible for such benefits. Blacks, as renters, would be more interested in enforcement of inspection codes to deal with substandard housing than white landlords who would have to comply.

The lower socioeconomic status of blacks in addition to shaping their goals also affects their ability to attain them. Generally, since blacks are more likely to be working class, they are less able to take time off to register and to vote. The lower median

Figure 2. Age and Sex Population Pyramid for Shell County.



income of the black community means less money available to contribute to political campaigns and less to spend on the media. Their limited opportunities for education and employment also encourage young blacks to leave the area. The results are shown in the age and sex population pyramid, Figure 2. The 20-24 yr. group shows a marked decrease for the black population, less among the whites ("A" on both pyramids). This means that the voting age population (VAP) will be a smaller % of their population for blacks than for whites, decreasing their political strength, and that the black VAP, relative to the white, has dropped further since the date of Census enumeration.

The election practices that were at issue in the lawsuit -- in our case, at-large elections rather than election by districts -- are the subtle contemporary counterparts to the old poll tax and "interpret the constitution" requirements. In conjunction with the age structure of the populations, the socioeconomic deterrents to black registration and voting, and the near-total white bloc voting, these election practices seriously hinder black political participation. This conclusion is the overall point of the expert testimony. In the general case, using census data (or other data, where available) can provide a context for the association found through correlation/regression, so that the court will understand not only that the association is significant, but what its importance is in the day-to-day functioning of the social structure.

REFERENCES

Alvarez, et al. (1979) Discrimination in Organizations. San Francisco: Jossey-Bass. Sample researches by social scientists studying discrimination in promotion, recruitment, and decision-making.

Anderson, T. and Zelditch, M. (1968) A Basic Course in Statistics. New York: Holt. Concise clear chapters.

Beere, C. (1979) Women and Women's Issues: A Handbook of Tests and Measures. San Francisco: Jossey-Bass. Of obvious relevance to possible sex discrimination cases (employment, curriculum, student admission, etc.); contains dozens of research instruments for measuring sex role preference, degree of work orientation, which jobs are perceived to be open to women, attitudes toward sex discrimination, male reactions to female competence, and many other topics. These instruments might carry more weight than a newly-devised questionnaire.

Becker, H. (1978) "The Measurement of Segregation," ERIC Document #ED171-825. Critiques the Taeuber Index of Dissimilarity.

Cohen, J. and P. (1975) Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. New York: Wiley. For a person needing more information than supplied in general statistics texts, this book offers a thorough presentation and argues that regression/correlation can be applied much more widely than it has been. Substantially more difficult than Loether/McTavish or Anderson/Zelditch.

Deutscher, I. (1973) What We Say/What We Do. Glenview: Scott, Foresman. To the social scientist, this book offers a sophisticated discussion of the pitfalls of relying on survey data. To the lawyer, the book suggests tough questions to put to survey researchers regarding the adequacy of their research designs and would therefore be useful in preparing for

cross-examination of one's own or an opposing expert witness.

Fiedler, J. (1978) Field Research. San Francisco: Jossey-Bass. Many (although by no means all) social science investigations of potential discrimination must be on-site; this book suggests ways of managing field research, particularly if the project must be large-scale.

Finkelstein, M. (1966) "The Application of Statistical Decision Theory to the Jury Discrimination Cases," Harvard Law Review, 80, 338-376. Argues for the usefulness of inferential statistics ("statistical decision theory" in his words) in various cases and explains clearly the test used. Since 1966, statistical analysis has come to play a much larger role in the courtroom, so some of his pleading for it strikes today's reader as antique.

---- (1978) Quantitative Methods in Law. New York: Free Press. The first chapter discusses the increasing use of statistics in legal cases; other useful chapters cover jury discrimination, voting (regarding one man/one vote reapportionment), and "regression models in administrative proceedings."

Freedman, D., et al. (1978) Statistics. New York: Norton. Massive thorough reference work with good emphasis on which test to use for which purpose.

Hallock, M. (1977) "The Numbers Game -- The Use and Misuse of Statistics in Civil Rights Litigation," Villanova Law Review, 23 #5, 5-34. Good discussion of inferential statistics and levels of significance, statistically and judicially considered. Treats jury and employment discrimination, demographic data, regression, and chi-square.

Hauser, P., ed. (1964) Handbook for Social Research in Urban Areas. Paris: UNESCO. A good introduction to doing research based on existing data (from the Census and other sources). The question of unit of analysis is well-discussed. Probable data sources are described. A good introduction to demographic analysis is provided by chapter 6. The book is a good starting reference for anyone involved with school desegration, municipal services discrimination, or other research or litigation involving urban areas.

Koosis, D. (1972) Statistics. New York: Wiley. Offers excellent step-by-step exercises to teach the use of some statistical techniques, including frequency distributions, sampling, and difference between two means.

Langbein, L. and Lichtman, A. (1978) Ecological Inference. Beverly Hills: Sage. Like others in the Sage series, this booklet is hard going sometimes, but it offers an authoritative and compact treatment of ecological regression, relevant to voting analyses and other areal topics.

LCCRUL (no date) A Community Guide to the Equalization of Municipal Services. Washington: Lawyers' Committee for Civil Rights Under Law. How to conduct a community survey to determine if street paving, parks, garbage collection, etc., are provided on a nondiscriminatory basis.

Lehmann, E. (1975) Nonparametrics. San Francisco: Holden-Day. This is an update of Siegel (q.v.) and is thorough and useful, but does not have Siegel's unique clarity and utility for the neophyte.

Loether, H. and McTavish, D. (1974) Descriptive Statistics for Sociologists. Inferential Statistics for Sociologists. Boston: Allyn and Bacon. These books are a little more difficult than Anderson/Zelditch but are more complete. Descriptive Statistics includes a useful "how-to" section on graphic presentation. The clear delineation between descriptive and inferential statistics helps avoid confusion.

Meyer, P. (1973) Precision Journalism. Bloomington: Indiana University Press. This sympathetic account of a journalist's encounter with social science and statistics makes good reading for the person who needs to be able to use ideas from those fields but has been put off by them in the past.

Miller, D. (1977) Handbook of Research Design and Social Measurement. New York: McKay. A compendium of books, scales, and other aspects of research design, helpfully arranged and annotated.

Nie, N., et al. (1975) Statistical Package for the Social Sciences, New York: McGraw-Hill. This manual supplies computer programs to calculate statistics commonly used in social science research.

Payne, S. (1951) The Art of Asking Questions. Princeton: Princeton University Press. This old book stays in print because it helps anyone making up a survey ensure that the questions are clear, the answers meaningful.

Raj, D. (1972) The Design of Sample Surveys. New York: McGraw-Hill. Useful in addition to Sudman because it gives specific surveys in its later pages, with particularly helpful chapters on demographic and employment surveys.

Reynolds, H. T. (1977) Analysis of Nominal Data. Beverly Hills: Sage. Sophisticated but compact presentation of chi-square and other ways to analyze nonparametric variables. Many analyses, including chi-square, usually thought of as measures of association, really amount to significance tests (part of inferential statistics); some useful but uncommon true measures of association are suggested as alternatives.

Siegel, S. (1956) Nonparametric Statistics. New York: McGraw-Hill. The classic source for many simple, yet powerful, statistical tests, including the Sign Test, Chi-square, Runs Test, Wilcoxon Signed-Ranks Test, and Spearman Correlation Coefficient. It is also clearly written; each section concludes with a crystalline "summary of procedure."

Slonim, M. (1966) Sampling in a Nutshell. New York: Simon and Schuster. A good simple introduction to sampling.

Smith, H. (1975) Strategies of Social Research. Englewood Cliffs, N. J.: Prentice-Hall. In short readable chapters this book explains each part of the research process from ethics to interpretation. Written for social scientists, but the lawyer or other nonscientist will not find the language difficult, only unfamiliar.

Sorensen, A., et al. (1975) "Indexes of Racial Residential Segregation for 109 Cities in the United States," Sociological Focus, 8, 125-142. Provides current indices of residential segregation.

Sudman, S. (1976) Applied Sampling. New York: Academic Press. A thorough treatment of sampling, with full discussion of small-scale sampling and costs, how large the sample must be, and different ways of taking it. Written for social scientists.

Taeuber, K. and A. (1965) Negroes in Cities. Chicago: Aldine. Analyzes residential segregation; proposes the Index of Dissimilarity.

Taeuber, K. and Wilson, F. (1979) Project Report #1: Analysis of Trends in School Segregation (draft). Madison: Institute for Research on Poverty. Provides current indices of school segregation.

38 Social Scientists (1979) "School Desegregation and Residential Segregation," Transaction/

Society, 16 #5, 70-76. Discusses the complex interrelationship between these two forms of segregation.

Underwood, B. (1979) "Law and the Crystal Ball: Predicting Behavior with Statistical Inference and Individualized Judgment," Yale Law Journal, 88, 1408-48. Discusses how to challenge selection systems using predictive criteria, such as aptitude tests or personal characteristics used predictively, as happens in school admissions, credit risk evaluation, likelihood of recidivism, etc.

Verba, S. and Nie, N. (1972) Participation in America. New York: Harper. Useful introduction to political science generalizations regarding voting behavior.

Webb, E., et al. (1966) Unobtrusive Measures. Chicago: Rand McNally. Suggests ways of getting data other than through surveys. The first chapter is marred by sociological jargon, but later chapters can spur the imaginative mind to develop sources of data that are particularly authoritative because they are "physical" rather than created through opinions voiced in interviews or questionnaires.

Weeks, J. (1978) Population. Belmont: Wadsworth. Provides a good basic understanding of demography. Part II explains the three basic population processes -- fertility, mortality, and migration; Part III explains the age and sex structure and the impact of the processes upon it. Clearly laid out, but not a "how-to" book and does not teach specific demographic methods.

Zelditch, M. (1959) A Basic Course in Sociological Statistics. New York: Holt. Even simpler and clearer than Anderson and Zelditch (q.v.).

Bibliographic note: In addition to the works listed above, three publisher's series are of particular importance regarding applications of methodology and statistics. Many of the works in these series are on the "cutting edge" of the discipline, and some are written so as to be unintelligible to the average social scientist, let alone the average nonscientist, which restricts their usefulness. Nonetheless, potential social science experts should be aware of them:

Quantitative Applications in the Social Sciences (series), Beverly Hills: Sage. See also other Sage books in evaluation research, methodology, and statistics.

Sociological Methodology (annual), San Francisco: Jossey-Bass. See also other Jossey-Bass books in methodology and statistics.

Wiley books in methodology and statistics, New York: Wiley.

CONTRIBUTORS

Kathleen H. Courtney, a former Center intern, brings a background in sociology and law to the CLEARINGHOUSE. Her B.A. was earned at the State University of New York at Plattsburgh in sociology; she is now completing her J.D. at Catholic University.

James W. Loewen edits the CLEARINGHOUSE and is Director of Research for the Center. He has testified in civil rights cases involving voting rights, jury discrimination, unequal housing assessments, and several other issues, and is on leave from the University of Vermont, where he is Associate Professor of Sociology.

William L. Taylor, Director of the Center, is a civil rights lawyer who has used social science evidence in several cases. As Staff Director of the U. S. Commission on Civil Rights in the 1960's he supervised several inter-disciplinary studies, among them Racial Isolation in the Public Schools.

COMMUNICATIONS TO THE CLEARINGHOUSE

The Spring, 1979, issue, "School and Residential Desegregation," continues to draw requests from persons who wish to use it in school systems and communities undergoing desegregation. The education coordinator of the human relations commission of a large city under court-ordered metropolitan desegregation writes:

> We are working now on a proposal which would modify our desegregation plan to require more busing of white children. Since the issue is a hotly political one, I could use some help. I think it might be advantageous to give our present school board members a copy of your Spring, 1979, CLEARINGHOUSE. Could you send me ten copies of that issue as soon as possible? I would be most grateful.

We did, and were happy to know that the issue was deemed useful. Another letter requested ten copies of our "Testing and Citizenship" double issue, Autumn-Winter, 1978, and said:

> Would you please rush this order. We need them to distribute to key schools before testing considerations for this year are complete.

CLEARINGHOUSE subscriptions have quadrupled in the last year. Many of our newer subscribers have requested back issues so they can have a complete file. The following issues were published:

Vol. I, #1, #2, #3, #4.
Vol. II, #1, #2.
Vol. III, #1.
Vol. IV, #1.
Vol. V, #1, #2.
Vol. VI, #1-2 (double issue). (Sold out; available in Xeroxed form only.)
Vol. VII, #1, #2, #3-4 (double issue).

We can supply back issues for $2 each ($4 for double issues). All fourteen issues published prior to this one are also available to any subscriber who prepays in the amount of $20 -- a discount of almost 30%.

SEEKING SOCIAL SCIENCE

The CLEARINGHOUSE wants to hear from social scientists or lawyers who have used specific research or statistical techniques in the court-room. Please write the Editor.

For the next issue, "Health Care and Minorities," we seek compact articles with national policy implications. Send two copies and a vita to the Editor. (The vita is not examined until after the article is reviewed; it helps us suggest directions for modification, if appropriate.) Upcoming issues will treat cultural pluralism, housing, and other topics. We solicit your articles or suggestions as to potential contributors.

The CLEARINGHOUSE disseminates current research in areas of civil rights and minority concern to people involved in influencing or making public policy. Authors are invited to submit papers summarizing findings with policy implications. Send articles or inquiries to James W. Loewen. Address subscription inquiries to "Editor".

The Center for National Policy Review is a group of lawyers and social scientists engaged in research and legal action on behalf of civil rights groups and in the interests of minorities. We monitor federal programs to help assure that legislative advances in civil rights are translated into policies that stem the tide of racial polarization.

William L. Taylor, Director
(Education and Title VI)

Roger S. Kuhn, Co-Director
(Credit Discrimination)

Glenda Sloane, Housing and Community Development Director

Garland Pinkston, Jr., Managing Attorney
(Employment and Minority Business)

Morton Sklar, Employment Project Director
(Federal grants, Employment)

James W. Loewen, Director of Research
(Desegregation, Cultural Pluralism)

James S. Byrne, Editor, JOBS WATCH
Erica Kearse, Editorial Assistant

Mary Tipton, Administrative Secretary
Sr. M. Adrienne, O.S.B., Admin. Secretary

Copyright, 1980, Center for National Policy Review. Reproduction permitted with prior notice.

Clearinghouse for Civil Rights Research

Center for National Policy Review
Catholic University School of Law
Washington, D. C. 20064
(202) 832-8525

Nonprofit Organization
U.S. Postage
Paid
Washington, D.C.
Permit No. 711

ERIC Clearinghouse on U. I
Teachers College
Columbia Univ., P.O.Box 4(
New York, NY 10027

Return postage guaranteed
Address correction requested